TECHNICAL RESEARCH REPORT

The Dynamics of Cross-Coupled, Self-Referential Linear Systems

by Mark Fleischer

CSHCN TR 2002-37
(ISR TR 2002-72)

THE DYNAMICS OF CROSS-COUPLED, SELF-REFERENTIAL LINEAR SYSTEMS

MARK A. FLEISCHER*

Abstract. This article presents the central ideas behind *cross-coupled, self-referential* linear systems—dual systems where each system in the pair provides the other system with reference or control signals using a form of state feedback. Such systems are ubiquitous in nature, the most noteworthy being the mammalian brain. Although complex systems and feedback mechanisms have several decades worth of literature, the fundamental aspects of simple cross-coupled linear systems have apparently not been fully explored or articulated. This surprising fact and the simplicity of the concepts involved provide a backdrop in which the fundamental nature of cross-coupled systems are investigated by examination of linear iterative maps. The *cross-coupling effect* in iterative maps is shown to reduce the magnitude of the eigenvalues of the linear system when the two inputs to the system are unequal. Applications to the solution of linear systems are also presented and shown to enlarge the applicability of the Gauss-Seidel iterative method. Self-similarity and scaling properties are also examined in which cross-coupled systems are cross-coupled. The eigenvalues in such systems have multiplicities described by Pascal's Triangle. Future research areas such as neural networks, control systems, and Markov Decision Processes are discussed including ideas on how such cross-coupled systems can serve as a model for autonomous control systems and even for human consciousness.

Key words. Dynamical systems, iterative maps, autonomous systems, feedback control, state feedback, optimal control, cybernetics, consciousness, cognitive processes.

AMS subject classifications. 26A18, 37C25, 37C45, 37C75, 39B05, 47H10

1. Introduction. In the realm of the arts and sciences, perspective is often the key ingredient in the achievement of new insights and knowledge. Norbert Wiener alluded to this fact when he said that

> . . . the most fruitful areas for the growth of the sciences were those which had been neglected as a no-man's land between the various established fields. [15]

This "no-man's land" is often a fruitful area because of the different perspectives a scientist brings in confronting a problem or developing an idea outside of his or her domain. This article is motivated by a perspective rooted in the subject of autonomous systems, artificial intelligence and control systems—in short, the field named by Wiener as *cybernetics* which comes from the Greek word for *steersman*. It is from this perspective that the idea of *how* to cross-couple *linear* systems was incubated. The reader will discover that it is just a bit ironic that this perspective leads to simple models that mathematically capture the value of utilizing different perspectives.

So how does an autonomous system take advantage of different perspectives and different system states so as to make better decisions? The motivating question, the heart of the matter is, how does an autonomous system control *itself* in a randomly changing environment? One possible answer proposed in this article is that two *cross-coupled* systems control each other and, in so doing, the aggregate system controls itself. This article explores this notion of cross-coupling by examining *pairs* of simple, linear iterative maps or dynamical systems.

*Institute for Systems Research, University of Maryland, College Park, Maryland 20742, (301) 405-2958 (`mfleisch@isr.umd.edu`)

Given that this article is about cross-coupling systems, just what is *cross-coupling*? Federal Standard 1037C defines cross-coupling as follows:

> *The coupling of a signal from one channel, circuit, or conductor to another, where it is usually considered to be an undesired signal.* [4]

This article extends this definition to include dynamical systems and modifies it somewhat by showing how cross-coupling can be a *desirable* feature in these systems. The goal here is to articulate a new paradigm for developing and describing whole systems based on pairs of identical (or similar) systems where each system of the pair provides the other with reference or control signals. In control theory parlance, each system provides *the other* system with a form of *state feedback*. Such systems therefore can be defined as *cross-coupled* and, because they are functionally dependent on each other and similar or identical, such system pairs are also *self-referential*. The cross-coupling thereby raises the complexity of the entire system—a complexity worth examining.

The motivating issues and ideas behind this article are more fully explored in Section 2. Simple linear dynamical systems or iterative maps are described in Section 3. This involves two approaches: 1) where output values are fed back to modify inputs; and 2) where the inputs are cross-coupled so that they have the same time index. Theorems are presented showing the relationships between input vectors, eigenvalues and eigenvectors and show how different inputs into a cross-coupled linear system can improve the convergence rate of the system to its fixed points.

Section 4 describes how cross-coupled systems can be scaled up to reflect self-similarity. The idea of cross-coupling cross-coupled systems is therefore examined. Patterns emerge with respect to the spectrum of eigenvalues of such systems—the multiplicity of eigenvalues is given by Pascal's Triangle. Potential applications of the theory and ideas for further research are discussed in Section 5. Finally, Section 6 summarizes the material and provides some concluding remarks.

2. Background and Motivation. Science often benefits by noting the connections between the established fields alluded to earlier in the quote from Wiener, but also from understanding the motivating concepts and domains of new ideas. To that end, a brief discourse on the motivating ideas behind cross-coupling is warranted.

The basic notion of cross-coupled systems comes from observations of nature. Nature and the mechanisms of evolution have provided us with a vast library of problem solving techniques, examples and illustrations. New ones are constantly being discovered—we only have to appreciate their simplicity to realize their power. With the rise in computing power and the development of new areas of mathematics such as chaos theory, it is now possible to more fully explore ideas and examples from nature. This ultimately has lead to ambitious efforts to understand the Holy Grail of complex systems—the human brain. In the last 50 plus years since the ideas of McCullogh and Pitts' perceptron [11, 14] first hinted at a paradigm for investigating and understanding the human brain, some ground has been gained towards understanding its fundamental nature and even that of human consciousness (see [5, 10]).

In fact, it was the structure of the brain and earlier work in parallel computing schemes for optimization that provoked the ideas behind cross coupling systems. The ideas behind *cybernetic optimization by simulated annealing* (COSA) are based on creating a feedback mechanism to form a self-referencing system [7, 9].

As the reader will no doubt discover, the conceptual foundation of cross-coupled systems is simple yet compelling and described in abstract terms below along with some discussion of how nature has evolved such systems in the human brain.

2.1. Simplicity and Beauty in Nature. Consider a pair of similar or identical systems where one system, denoted as $\mathbf{S_1}$, provides the other system of the pair, $\mathbf{S_2}$, with a better measure of the limiting or desired system state. In such a system, this state information can be used to improve the system state or performance measure (*e.g.,* faster convergence) of $\mathbf{S_2}$ which in turn can lead to a better state/performance in $\mathbf{S_1}$ and so on. Such a dynamic may therefore give rise to a multiplier effect or "hall-of-mirrors" dynamic. Figure 2.1 depicts this concept using two identical cross-coupled, self-referential systems. Cross-coupling may, therefore, provide systems with

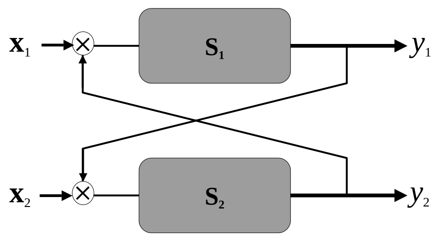

FIG. 2.1. *A Simple Cross-Coupled, Self-Referential System*

advantageous dynamics that stem not from the apparent redundancies in the pair, but from the cross-coupling itself.

This is most easily appreciated in the context of iterative maps. Let $y = f(\mathbf{x})$ be some input/output function for input vector $\mathbf{x}$, with scalar output y. In the associated cross-coupled system (2.1), the function is modified and its argument list expanded so that two functions $f_1(\cdot) \equiv f_2(\cdot)$ and feedback is used to modify the outputs. Note that the feedback to $f_1(\cdot)$ comes not only from that function's output, but also from the other identical (or possibly similar) system $f_2(\cdot)$. Thus,

$$\begin{array}{rcl} y_1 & = & f(\mathbf{x}_1, y_1, y_2) \\ y_2 & = & f(\mathbf{x}_2, y_2, y_1) \end{array} \qquad (2.1)$$

In this cross-coupled system, y_1 is not only a function of itself, *i.e.,* y_1, but also of y_2 *which is also* a function of y_1. Consequently, the output y_1 is a more complex function of the inputs and feedback values. In some sense, it is a function of itself after *two* processes defined by the function $f(\cdot)$. The cross-coupling therefore establishes a higher degree of self-referencing and feedback that can markedly affect system dynamics and performance.

This is perhaps why such systems are so prevalent in nature. In fact, cross-coupled systems are ubiquitous: the mammalian brain is perhaps the most obvious example of a cross-coupled system. It seems to have evolved as a simple mechanism for increasing the survival value of animals. Its beneficial effects are clearly demonstrated in stereoscopic vision whereby pairs of optical systems, slightly displaced from one another, provide the capability of depth perception and gradient estimation. The displacement, which gives rise to differences in system states is, of course, crucial for the advantage of stereoscopic vision to be realized. The modelling and analysis that follows seems to account, to some degree, for this fact.

Cross-coupling may also constitute some clever control engineering. For control systems to function effectively, reference or control signals must somehow be made available to the system. Reference signals allow a complex system to gauge and control its behavior. But in a complex, changing, and uncertain environment such as the natural environment, from where are such reference signals to come?

Nature may have solved this problem by a clever trick that has both biological and mathematical beauty and simplicity—providing two brains, or more accurately, two hemispheres in a single brain where each hemisphere provides at least some of the reference and/or control signals to the other hemisphere. The system as a whole therefore can *fake* the acquisition of reference signals by virtue of this cross-coupling.

One certainly can marvel at the notion that humans have two hemispheres because the cross-coupling phenomenon is advantageous to survival. But how is it that nature could have evolved such a clever mechanism? The answer to this question may come from understanding the embryonic phases of fetal development known as *gastrulation* followed by *longitudinal folding* [12].

During gastrulation, the cell structure of the human embryo begins to differentiate into three distinct layers and flatten into a disc shape. After the three layers are formed, they begin to grow at different rates which initiates the process of longitudinal folding where the embryonic disc folds in on itself. The edges of the folded disk knit together to form the neural folds and tubes, *i.e.,* the spine and brain-stem [12]. This folding operation also creates the apparent symmetry in the human body including the symmetry in brain structure (of course, this symmetry is not absolute). Nature thus seems to have evolved a *conceptually simple mechanism* to create a cross-coupled system. As such, it bears further investigation. Inquiry in this particular phenomenon has, however, been rather sparse.

2.2. The Historical Record on Cross-Coupled Systems. As indicated earlier, cross-coupled systems abound not only in nature in the form of brains, but also in our technology. Yet, surprisingly, there does not seem to be any analysis in the literature on cross-coupled discrete linear dynamical systems or iterative maps. This surprising state of affairs may stem from a long-standing bias and inclination on the part of engineers and scientists to find ways to *de*-couple systems to make them more amenable to analysis. The definition of *cross-coupling* in the Federal Standard seems to bear this out.

When one examines the literature in complex systems such as neural nets, dynamical systems, control systems and others, no mention is made of this concept. Even for the most interesting case, the human brain, the cross-coupling phenomenon is not mentioned *per se*. To be sure, texts on the human brain often mention the *corpus callosum*, the bundle of nerve tissue that connects the two hemispheres of the brain and establishes the cross-coupling in the human brain, but the discussions usually are limited only to the most medical of contexts. Descriptions of experiments with split brains (where the *corpus callosum* is severed), the psychological and cognitive changes that occur in patients with split brains are stated with pathological terminology, but never in positive terms such as how cross-coupling might be *advantageous* for cognition, control, or consciousness. More often, the explanation of why two hemispheres evolved is described in terms of redundancy and efficiencies such as how the left hemisphere controls the right side of the body, etc. (see *e.g.,* [2, 15]). It is entirely possible, however, that the two hemispheres of the brain evolved because of how a cross-coupling dynamic significantly enhances survival-value, a dynamic that may constitute the most fundamental aspect of human consciousness.

Figure 2.1 should look quite familiar to electrical engineers and computer scientists as it resembles the *flip-flop*, a cross coupled system based on logic circuits. The elements of memory and switching circuits are all based on the flip-flop, yet their nature is couched in terms of stability, their utility for holding a state, or as a bounceless switch and so forth but not as dynamical systems or iterative maps. Indeed, even in the most excellent texts on dynamical systems, the nature of feedback mechanisms and such are exhaustively described except for cross-coupled, self-referencing systems! See for example [1, 2] (cf. [2] for a description of self-referencing systems). Many articles describe cross-coupled *control systems*, but in very classical and limited terms. For example, [16] describes cross-coupled motion control systems, but their description is limited to the specific purpose of motion control systems and does not consider the the fundamental effects of coupling on eigenvalues, etc.

The goal of this paper, therefore, is to highlight what seems to be a phenomenon that deserves to be examined mathematically. If we consider the possibility that many of the reference or control signals in each hemisphere of the human brain comes from the other hemisphere, then we need to examine what, if any, beneficial effects such cross-coupling enables. To do this in a way that highlights *fundamental* properties, the most simple of cross-coupled systems should be examined, at least initially. The goal here is rather modest and seeks to determine any interesting and fundamental aspects of cross-coupled systems. In this vein, the following exploration uses simple, linear equations and systems of equations as dynamical systems.

3. Cross-Coupled Linear Systems. This section examines two types of simple linear systems: one where output values (*i.e.,* state variable with the next time index) are used as reference signals; and 2) where each system in the pair utilize the other system's inputs as a reference signal (*i.e.,* with the same time index as the input). For lack of better terminology, the former will be denoted as an *output-based* cross-coupled system and the latter as an *input-based* cross-coupled system. First, let us consider the basic mathematics associated with cross-coupled linear equations as iterative maps.

3.1. Cross-Coupled Linear Equations. Suppose one desires to calculate a fixed point $\mathbf{x}^* = f(\mathbf{x}^*)$ in the dynamical system

$$x^{[k+1]} = ax^{[k]} + c \tag{3.1}$$

where time indices are used to illustrate the iterative dynamics. If $|a| < 1$, any initial value of x, *i.e.,* $x^{[0]}$ converges to $x^* = \frac{c}{1-a}$. The error

$$\begin{aligned} e^{[k+1]} &= x^{[k+1]} - x^* = ax^{[k]} + c - (ax^* + c) \\ &= a(x^{[k]} - x^*) = ae^{[k]} \end{aligned}$$

therefore converges geometrically to zero at the rate of a per iteration.

3.1.1. Using Perfect Information for Reference Signals. Now suppose we desire that the sequence of iterations converge to the solution x^* at a faster rate. If we have *perfect information* as to the ultimate solution x^*, we could incorporate a feedback mechanism so that a fraction of any error with respect to x^* in an iterate is subtracted from the iterate in the next iteration. Consider the following two parallel, identical and *independent* systems (because they are independent, only one equation is necessary to show how perfect information can be utilized. Presenting it this way, however, facilitates the later discussion):

$$
\begin{array}{rcl}
x_1^{[k+1]} & = & a(x_1^{[k]} - \omega[x_1^{[k+1]} - x^*]) + c \\
x_2^{[k+1]} & = & a(x_2^{[k]} - \omega[x_2^{[k+1]} - x^*]) + c.
\end{array} \tag{3.2}
$$

where a fraction ω of the error in the output $x^{[k+1]} - x^*$ is subtracted from the input in each equation. In matrix-vector notation this becomes

$$
\mathbf{x}^{[k+1]} = \mathbf{A}\left(\mathbf{x}^{[k]} - \omega\left[\mathbf{x}^{[k+1]} - \mathbf{x}^*\right]\right) + \mathbf{c} \tag{3.3}
$$

where

$$
\mathbf{A} = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix},\ \mathbf{x}^{[k+1]} = \begin{pmatrix} x_1^{[k+1]} \\ x_2^{[k+1]} \end{pmatrix},\ \mathbf{x}^* = \begin{pmatrix} x^* \\ x^* \end{pmatrix} \text{ and } \mathbf{c} = \begin{pmatrix} c \\ c \end{pmatrix}.
$$

Note that matrix $\mathbf{A}$ is symmetric. Solving the system in (3.3) to render it into feedforward form, gives

$$
\mathbf{x}^{[k+1]} = (\mathbf{I} + \omega\mathbf{A})^{-1}(\mathbf{A}\mathbf{x}^{[k]} + \omega\mathbf{A}\mathbf{x}^* + \mathbf{c}). \tag{3.4}
$$

In this case, the coefficient matrix in (3.4) after some algebra is

$$
(\mathbf{I} + \omega\mathbf{A})^{-1}\mathbf{A} = \begin{pmatrix} \frac{a}{1+\omega a} & 0 \\ 0 & \frac{a}{1+\omega a} \end{pmatrix}
$$

and now the system converges at the rate of $\frac{a}{1+\omega a}$. If ω has the same sign as a, then the use of perfect information increases the convergence rate by a factor $1/(1+\omega a) > 0$.

In this dynamical system, the fixed point x^* in effect provides a *reference* or *control signal* for each equation. Obviously, knowledge of the fixed point allows the error signal to be accurately gauged and measured. Perfect information allows the error to be perfectly determined. Suppose however that we do not have perfect information and consequently do not know the value of the fixed point solution. Is it possible to improve the convergence rate by approximating the reference value in one system by using information from the other identical dynamical system?

3.1.2. Using Imperfect Information for Reference Signals. Equation (3.5) shows how two *cross-coupled* systems can supply an approximation of these reference signals to each other. Consider

$$
\begin{array}{rcl}
x_1^{[k+1]} & = & a(x_1^{[k]} - \omega[x_1^{[k+1]} - x_2^{[k+1]}]) + c \\
x_2^{[k+1]} & = & a(x_2^{[k]} - \omega[x_2^{[k+1]} - x_1^{[k+1]}]) + c
\end{array} \tag{3.5}
$$

where the output values of each equation replace x^* and serve as the reference value in the other equation. Thus, $x_2^{[k+1]}$ is substituted for x^* in the first equation and $x_1^{[k+1]}$ is substituted for x^* in the second equation. These two equations constitute a cross-coupled system. The matrix form is

$$
\mathbf{x}^{[k+1]} = \mathbf{A}\left(\mathbf{x}^{[k]} - \omega\left[\mathbf{x}^{[k+1]} - \mathbf{P}\mathbf{x}^{[k+1]}\right]\right) + \mathbf{c} \tag{3.6}
$$

where the identity

$$
\begin{pmatrix} x_2^{[k+1]} \\ x_1^{[k+1]} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1^{[k+1]} \\ x_2^{[k+1]} \end{pmatrix}
$$

is used to effect the cross-coupling using the permutation matrix $\mathbf{P} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ in (3.6). In feedforward form, this becomes

$$\mathbf{x}^{[k+1]} = [\mathbf{I} + \omega\mathbf{A} - \omega\mathbf{AP}]^{-1} (\mathbf{Ax}^{[k]} + \mathbf{c}). \tag{3.7}$$

The coefficient matrix in (3.7)

$$\begin{aligned} \mathbf{C} = [\mathbf{I} + \omega\mathbf{A} - \omega\mathbf{AP}]^{-1} &= \begin{pmatrix} 1+\omega a & -\omega a \\ -\omega a & 1+\omega a \end{pmatrix}^{-1} \\ &= \frac{1}{1+2\omega a} \begin{pmatrix} 1+\omega a & \omega a \\ \omega a & 1+\omega a \end{pmatrix} \end{aligned} \tag{3.8}$$

will be denoted as the *cross-coupling matrix*. To assess the convergence rate of this system, the eigenvalues of the matrix

$$\mathbf{CA} = \begin{pmatrix} \frac{a+\omega a^2}{1+2\omega a} & \frac{\omega a^2}{1+2\omega a} \\ \frac{\omega a^2}{1+2\omega a} & \frac{a+\omega a^2}{1+2\omega a} \end{pmatrix} \tag{3.9}$$

must be determined. Saving the proofs for the more general case below, the eigenvalues of (3.9) can be determined by using the characteristic equation

$$\left(\frac{a+\omega a^2}{1+2\omega a} - \lambda\right)^2 - \left(\frac{\omega a^2}{1+2\omega a}\right)^2 = 0 \tag{3.10}$$

and applying the quadratic formula. In this case, the set of eigenvalues are $\{\frac{a}{1+2\omega a}, a\}$ and the corresponding eigenvectors are

$$\begin{pmatrix} 1 \\ -1 \end{pmatrix} \text{ and } \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

respectively, which form an orthogonal eigenspace (this follows directly from the fact that the cross-coupling matrix is symmetric [6, 13]).

This permits a number of interesting observations. When an input vector has identical components, *i.e.,* is some multiple of the eigenvector $[1,1]^T$, the corresponding eigenvalue is a. Under these conditions, the feedback or output value $x_1^{[k+1]}$ in the first equation in (3.5) is identical to the reference value $x_2^{[k+1]}$, hence the error component is zero and no useful feedback information can be utilized. In this case, the convergence rate is a, identical to the original system in (3.1). For any other set of input values that differ, *i.e.,* where $x_1 \neq x_2$, the convergence rate is faster than a (*i.e.,* some linear combination of $\frac{a}{1+2\omega a}$ and a). What is remarkable and somewhat surprising, however, is that if the input values differ by a sign change where $x_1 = -x_2$, the cross-coupled system converges faster than with perfect information! This is because the denominator associated with the eigenvalues of the cross-coupled system is $1 + 2\omega a$ versus $1 + \omega a$, the denominator in the eigenvalue of the uncrossed coupled system in (3.4).

These results suggest how cross-coupling can affect dynamical systems in a desirable way and show how it is possible for two identical systems *in different states* to provide useful reference signals to each other. The computation performed by the first system provides useful information to the second which then provides even more useful information to the first and so on. This gives rise to the multiplier effect alluded to in Section 1 as evidenced by the eigenvalues of the matrix in (3.9). Thus, cross-coupling two identical systems can improve the convergence rate of an equation to its fixed-point over that of a system that does not use this form of state feedback. In a sense, the cross-coupled system 'knows' itself, *i.e.,* its states, better than two independent systems.

3.2. Cross-Coupled *Output-Based* Systems. In this section a similar model as the one described earlier is presented in which output values of one process serve as reference signals for the other process. This model extends the simple cross-coupled pair of linear equations to cross-coupled pairs of linear systems of equations. This system will be denoted as an *output-based* system. Consider the linear system

$$\mathbf{x}^{[k+1]} = \mathbf{A}\mathbf{x}^{[k]} + \mathbf{c} \tag{3.11}$$

where $\mathbf{x}^{[k]}$ is some n-vector and $\mathbf{A}$ is an $n \times n$ non-singular matrix. If an identical system were used to provide a reference or control value to the first and vice versa, then these two systems constitute a cross-coupled linear *system* as in

$$\begin{aligned} \mathbf{x}_1^{[k+1]} &= \mathbf{A}(\mathbf{x}_1^{[k]} - \omega[\mathbf{x}_1^{[k+1]} - \mathbf{x}_2^{[k+1]}]) + \mathbf{c} \\ \mathbf{x}_2^{[k+1]} &= \mathbf{A}(\mathbf{x}_2^{[k]} - \omega[\mathbf{x}_2^{[k+1]} - \mathbf{x}_1^{[k+1]}]) + \mathbf{c}. \end{aligned} \tag{3.12}$$

The following definitions and conventions will be used to denote various block matrices and vectors:

Definition 3.1. *Define the block matrices*

$$\hat{\mathbf{A}} = \left(\begin{array}{c|c} \mathbf{A} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{A} \end{array}\right) \text{ and block permutation matrix } \hat{\mathbf{P}} = \left(\begin{array}{c|c} \mathbf{0} & \mathbf{I} \\ \hline \mathbf{I} & \mathbf{0} \end{array}\right).$$

Define generic block vectors with possibly distinct vector components by

$$\hat{\mathbf{x}}^{[k]} = \begin{bmatrix} \mathbf{x}_1^{[k]} \\ \mathbf{x}_2^{[k]} \end{bmatrix}$$

and block vectors with common components by

$$\breve{\mathbf{x}}^{[k]} = \begin{bmatrix} \mathbf{x}^{[k]} \\ \mathbf{x}^{[k]} \end{bmatrix} \text{ and } \widetilde{\mathbf{x}}^{[k]} = \begin{bmatrix} \mathbf{x}^{[k]} \\ -\mathbf{x}^{[k]} \end{bmatrix}$$

where the $\smile$ *accent indicates the vector components are equal and the* $\sim$ *accent indicates the vector components are additive inverses.*

Using these definitions and conventions and rewriting (3.12) in matrix-vector form yields

$$\hat{\mathbf{x}}^{[k+1]} = \hat{\mathbf{A}}(\hat{\mathbf{x}}^{[k]} - \omega[\hat{\mathbf{x}}^{[k+1]} - \hat{\mathbf{P}}\hat{\mathbf{x}}^{[k+1]}]) + \breve{\mathbf{c}} \tag{3.13}$$

System (3.13) can be analyzed in the same fashion as was (3.5). In feedforward form, system (3.13) becomes

$$\hat{\mathbf{x}}^{[k+1]} = \left(\mathbf{I} + \omega\hat{\mathbf{A}} - \omega\hat{\mathbf{A}}\hat{\mathbf{P}}\right)^{-1}(\hat{\mathbf{A}}\hat{\mathbf{x}}^{[k]} + \breve{\mathbf{c}}). \tag{3.14}$$

The cross-coupling matrix for the linear system in (3.14) has the same form as in (3.7) only consists of block matrices. Thus,

$$\hat{\mathbf{C}} = \left(\mathbf{I} + \omega\hat{\mathbf{A}} - \omega\hat{\mathbf{A}}\hat{\mathbf{P}}\right)^{-1} = \begin{pmatrix} \mathbf{I}+\omega\mathbf{A} & -\omega\mathbf{A} \\ -\omega\mathbf{A} & \mathbf{I}+\omega\mathbf{A} \end{pmatrix}^{-1} \tag{3.15}$$

$$= \begin{pmatrix} (\mathbf{I}+2\omega\mathbf{A})^{-1}(\mathbf{I}+\omega\mathbf{A}) & \omega(\mathbf{I}+2\omega\mathbf{A})^{-1}\mathbf{A} \\ \omega(\mathbf{I}+2\omega\mathbf{A})^{-1}\mathbf{A} & (\mathbf{I}+2\omega\mathbf{A})^{-1}(\mathbf{I}+\omega\mathbf{A}) \end{pmatrix}$$

$$= \begin{pmatrix} (\mathbf{I}+2\omega\mathbf{A})^{-1} & \mathbf{0} \\ \mathbf{0} & (\mathbf{I}+2\omega\mathbf{A})^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{I}+\omega\mathbf{A} & \omega\mathbf{A} \\ \omega\mathbf{A} & \mathbf{I}+\omega\mathbf{A} \end{pmatrix}$$

$$= \hat{\mathbf{B}}\left(\mathbf{I} + \omega\hat{\mathbf{A}} + \omega\hat{\mathbf{A}}\hat{\mathbf{P}}\right) \tag{3.16}$$

where block matrix $\hat{\mathbf{B}}$ has the matrix $\mathbf{B} = (\mathbf{I} + 2\omega\mathbf{A})^{-1}$ on the diagonals and the matrices $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ are assumed to be non-singular. Note the similarity of (3.16) to the matrix in (3.8). Instead of being multiplied by the determinant, the matrix elements in (3.16) are each multiplied by an analog to the determinant, the matrix $(\mathbf{I}+2\omega\mathbf{A})^{-1}$. The reader can verify that (3.16) is in fact the inverse matrix of (3.15). Observe that if $\mathbf{A}$ is symmetric, the cross-coupling matrix in (3.16) must also be symmetric although the results below are not restricted to symmetric matrices. Note also that the block diagonal matrices $\hat{\mathbf{A}}, \hat{\mathbf{B}}$ commute.

The question arises as to how the cross-coupling matrix $\hat{\mathbf{C}}$ affects the dynamics in system (3.14). Before considering this, however, the mathematical form of a cross-coupled *input-based* system is examined. As will become apparent, these two forms have a number of mathematical similarities and performance characteristics.

3.3. Cross-Coupled *Input-Based* Systems. One of the draw backs of the output-based system is that output values lie on both sides of equation (3.12), *i.e.,* outputs of the linear system are used as inputs. This requires algebraic manipulation and matrix inversions to render it into feedforward form. While this approach is amenable to analysis for certain classes of linear systems, it may not provide a *realizable* system for many other functions or mappings that do no readily admit to inversion. Furthermore, the necessity of inverting the relevant matrices increases the computational burdens associated with using an output-based formulation.

This raises the question of whether it is possible to avoid using outputs, *i.e.,* iterates with time index $k+1$, for purposes of state feedback and still have the benefits of cross-coupling. In fact, the multiplier effect alluded to earlier is still present for input-based systems because the input values are merely the output values of the prior iteration. This section demonstrates that the cross-coupling of *inputs*, *i.e.,* iterates with time index k, yields similar benefits insofar as eigenvalues and eigenvectors are concerned. Indeed, the results associated with input-based systems are quite analogous to those obtained for output-based systems.

To that end, these results involve the following cross-coupled system, similar to (3.12) but using only input values, *i.e.,* values that have the same time index. Thus, the error signal is based on the difference between an input value and the reference signal, *i.e.,* the *input* of the other system. Thus,

$$\begin{aligned} \mathbf{x}_1^{[k+1]} &= \mathbf{A}(\mathbf{x}_1^{[k]} - \omega[\mathbf{x}_1^{[k]} - \mathbf{x}_2^{[k]}]) + \mathbf{c} \\ \mathbf{x}_2^{[k+1]} &= \mathbf{A}(\mathbf{x}_2^{[k]} - \omega[\mathbf{x}_2^{[k]} - \mathbf{x}_1^{[k]}]) + \mathbf{c}. \end{aligned} \tag{3.17}$$

Notice that (3.17) is identical to the system in (3.12) except for the time indices on the right-hand side of (3.17) which are all k. In matrix-vector form and after some simplification (without any matrix inversions), (3.17) becomes

$$\hat{\mathbf{x}}^{[k+1]} = \hat{\mathbf{A}}(\mathbf{I} - \omega\mathbf{I} + \omega\hat{\mathbf{P}})\hat{\mathbf{x}}^{[k]} + \check{\mathbf{c}} = [(1-\omega)\hat{\mathbf{A}} + \omega\hat{\mathbf{A}}\hat{\mathbf{P}}]\hat{\mathbf{x}}^{[k]} + \check{\mathbf{c}}. \tag{3.18}$$

In this case, the effect of the permutation matrix, which effectuates the cross-coupling, leads to a coefficient matrix that has convex combinations of the matrix $\mathbf{A}$ in each row and column. Thus,

$$\hat{\mathbf{A}}\hat{\mathbf{C}}' = (1-\omega)\hat{\mathbf{A}} + \omega\hat{\mathbf{A}}\hat{\mathbf{P}} = \begin{pmatrix} (1-\omega)\mathbf{A} & \omega\mathbf{A} \\ \omega\mathbf{A} & (1-\omega)\mathbf{A} \end{pmatrix} \tag{3.19}$$

where the matrix $\hat{\mathbf{C}}' = (1-\omega)\mathbf{I} + \omega\hat{\mathbf{P}}$. The $'$ is used here to distinguish $\hat{\mathbf{C}}'$ from the matrix $\hat{\mathbf{C}}$ associated with the output-based system.

In some ways, this matrix is more interesting than the one obtained in the output-based system. This matrix parametrically changes from an uncross-coupled system when $\omega = 0$, to a degenerate system when $\omega = \pm 1/2$ and to an equivalent uncrossed-coupled system with inputs switched when $\omega = 1$. As will become apparent from the analysis below, when $|\omega| < 1/2$, the system reduces the magnitude of the modified eigenvalues, and when $|\omega| > 1/2$ it increases them. Also, if the spectral radius of $\rho(\mathbf{A}) = 1$, varying ω can change the system dynamics from an attractor to a repeller or vice versa. Varying the value of ω also changes the matrix from positive to negative definiteness or vice versa. At this point, it is easy to see, and therefore stated without a formal proof, that

- if $\mathbf{A}$ is symmetric, then $\hat{\mathbf{A}}\hat{\mathbf{C}}'$ is symmetric.
- the sum of the elements in columns j and $n+j$ of $\hat{\mathbf{A}}\hat{\mathbf{C}}'$ equals the sum of the elements in column j of $\mathbf{A}$. Similarly, for rows.

Finally, notice that no matrix inversions are required to calculate $\hat{\mathbf{A}}\hat{\mathbf{C}}'$.

3.4. Analysis of Cross-Coupled Linear Systems. This section provides an analysis of the two cross-coupled models above. Theorem 3.2 describes the relationships among the eigenvalues and eigenvectors of the system in (3.11) and their cross-coupled counterparts (output-based and input-based) in systems (3.14) and (3.18).

THEOREM 3.2. *Let matrix $\mathbf{A}$ be a $n \times n$ matrix in the linear dynamical system $\mathbf{x}^{[k+1]} = \mathbf{A}\mathbf{x}^{[k]} + \mathbf{c}$. For the output-based system let $\omega > 0$ if $\mathbf{A}$ is positive definite and $\omega < 0$ if $\mathbf{A}$ is negative definite and for the input system let $\omega \in (-1/2, 1/2)$. Using the block matrix notation defined above, then the following statements are true:*

i–Output System *For each eigenvalue of $\mathbf{A}$,* i.e., *$\lambda_i(\mathbf{A})$, the system $\hat{\mathbf{C}}\hat{\mathbf{A}}$ has two eigenvalues, the original value $\lambda_i(\mathbf{A})$ and a* modified *eigenvalue $\frac{\lambda_i(\mathbf{A})}{1+2\omega\lambda_i(\mathbf{A})}$ where $\left|\frac{\lambda_i(\mathbf{A})}{1+2\omega\lambda_i(\mathbf{A})}\right| < |\lambda_i(\mathbf{A})|$.*

ii–Input System *For each eigenvalue of $\mathbf{A}$,* i.e., *$\lambda_i(\mathbf{A})$, the system $\hat{\mathbf{A}}\hat{\mathbf{C}}'$ has two eigenvalues, the original value $\lambda_i(\mathbf{A})$ and a* modified *eigenvalue $(1-2\omega)\lambda_i(\mathbf{A})$ where $|(1-2\omega)\lambda_i(\mathbf{A})| < |\lambda_i(\mathbf{A})|$.*

iii–Eigenvectors *For each eigenvector $\mathbf{v}_i$ of $\mathbf{A}$, the systems $\hat{\mathbf{C}}\hat{\mathbf{A}}$ and $\hat{\mathbf{A}}\hat{\mathbf{C}}'$ have two eigenvectors,*

$$\check{\mathbf{v}}_i = \begin{bmatrix} \mathbf{v}_i \\ \mathbf{v}_i \end{bmatrix} \text{ and } \tilde{\mathbf{v}}_i = \begin{bmatrix} \mathbf{v}_i \\ -\mathbf{v}_i \end{bmatrix}$$

corresponding to the original and modified eigenvalues, respectively.

Proof. **Statement i:** This proof is based on decomposing the matrix $\hat{\mathbf{C}}\hat{\mathbf{A}}$ into its three terms, determining the eigenvalues of each term, and taking direct sums [3] to determine the eigenvalues of $\hat{\mathbf{C}}\hat{\mathbf{A}}$. Therefore,

$$\begin{aligned}\hat{\mathbf{C}}\hat{\mathbf{A}} &= \hat{\mathbf{B}}\left(\mathbf{I} + \omega\hat{\mathbf{A}} + \omega\hat{\mathbf{A}}\hat{\mathbf{P}}\right)\hat{\mathbf{A}} \\ &= \hat{\mathbf{B}}\hat{\mathbf{A}} + \omega\hat{\mathbf{B}}\hat{\mathbf{A}}^2 + \omega\hat{\mathbf{B}}\hat{\mathbf{A}}\hat{\mathbf{P}}\hat{\mathbf{A}}. \end{aligned} \tag{3.20}$$

For convenience, the notation $\lambda(\cdot)$ is used here to denote a set of eigenvalues. Thus, an eigenvalue of $\mathbf{A}$, $\lambda_i(\mathbf{A}) \in \lambda(\mathbf{A})$. Therefore, let $\lambda(\hat{\mathbf{B}}\hat{\mathbf{A}})$ denote the set of eigenvalues associated with the matrix $\hat{\mathbf{B}}\hat{\mathbf{A}}$. The characteristic matrix is

$$\hat{\mathbf{B}}\hat{\mathbf{A}} - \lambda\mathbf{I} = \begin{pmatrix} \mathbf{BA} - \lambda\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{BA} - \lambda\mathbf{I} \end{pmatrix}$$

and the eigenvalues can be determined by solving the equation

$$|(\mathbf{BA} - \lambda\mathbf{I})|^2 = 0.$$

It is therefore sufficient to solve the equation $|\mathbf{BA} - \lambda\mathbf{I}| = 0$, the characteristic equation of matrix $\mathbf{BA}$ (this implies that each eigenvalue of $\hat{\mathbf{B}}\hat{\mathbf{A}}$ has a multiplicity of 2). Then,

$$\begin{aligned}\lambda(\hat{\mathbf{B}}\hat{\mathbf{A}}) = \lambda(\mathbf{BA}) &= \lambda\left((\mathbf{I} + 2\omega\mathbf{A})^{-1}\mathbf{A}\right) \\ &= \frac{1}{\lambda(\mathbf{A}^{-1}(\mathbf{I} + 2\omega\mathbf{A}))} \\ &= \frac{1}{\lambda(\mathbf{A}^{-1} + 2\omega\mathbf{I})} \end{aligned} \tag{3.21}$$

where the method of direct sums can also be applied to (3.21). Thus,

$$\lambda(\mathbf{A}^{-1} + 2\omega\mathbf{I}) = \frac{1}{\lambda(\mathbf{A})} + 2\omega = \frac{1}{\lambda(\mathbf{A})} + 2\omega\frac{\lambda(\mathbf{A})}{\lambda(\mathbf{A})}. \tag{3.22}$$

Substituting (3.22) into (3.21) yields the following equality for the first set of eigenvalues involved in (3.20):

$$\lambda(\hat{\mathbf{B}}\hat{\mathbf{A}}) = \frac{\lambda(\mathbf{A})}{1 + 2\omega\lambda(\mathbf{A})}. \tag{3.23}$$

Note that the eigenvalues of $\hat{\mathbf{B}}\hat{\mathbf{A}}$ are defined in terms of the eigenvalues of matrix $\mathbf{A}$.

The second set of eigenvalues of (3.20) is $\lambda(\omega\hat{\mathbf{B}}\hat{\mathbf{A}}^2)$. It is again sufficient for the same reasons of symmetry to determine the set $\lambda(\omega\mathbf{BA}^2)$. Thus,

$$\begin{aligned}\lambda(\omega\hat{\mathbf{B}}\hat{\mathbf{A}}^2) = \lambda(\omega\mathbf{BA}^2) &= \frac{\omega}{\lambda(\mathbf{A}^{-2}\mathbf{B}^{-1})} \\ &= \frac{\omega}{\lambda(\mathbf{A}^{-2}(\mathbf{I} + 2\omega\mathbf{A}))} \\ &= \frac{\omega}{\lambda(\mathbf{A}^{-2} + 2\omega\mathbf{A}^{-1})} \\ &= \frac{\omega}{\frac{1}{\lambda^2(\mathbf{A})} + \frac{2\omega}{\lambda(\mathbf{A})}} \\ &= \frac{\omega\lambda^2(\mathbf{A})}{1 + 2\omega\lambda(\mathbf{A})}. \end{aligned} \tag{3.24}$$

Finally, the set of eigenvalues of the third term in (3.20) must be determined. In this case, care must be taken to consider the effect of the permutation matrix $\hat{\mathbf{P}}$. For this third term, the characteristic matrix is

$$\omega\hat{\mathbf{B}}\hat{\mathbf{A}}\hat{\mathbf{P}}\hat{\mathbf{A}} - \lambda\mathbf{I} = \begin{pmatrix} -\lambda\mathbf{I} & \omega\mathbf{B}\mathbf{A}^2 \\ \omega\mathbf{B}\mathbf{A}^2 & -\lambda\mathbf{I} \end{pmatrix}.$$

Hence, the characteristic equation is

$$\left|(-\lambda\mathbf{I})^2 - (\omega\mathbf{B}\mathbf{A}^2)^2\right| = 0$$

or,

$$|(\lambda\mathbf{I} - \omega\mathbf{B}\mathbf{A}^2)(\lambda\mathbf{I} + \omega\mathbf{B}\mathbf{A}^2)| = 0.$$

Multiplying each factor by -1 yields,

$$\begin{aligned} |(\omega\mathbf{B}\mathbf{A}^2 - \lambda\mathbf{I})(-\omega\mathbf{B}\mathbf{A}^2 - \lambda\mathbf{I})| &= 0 \qquad \text{hence} \\ |\omega\mathbf{B}\mathbf{A}^2 - \lambda\mathbf{I}||-\omega\mathbf{B}\mathbf{A}^2 - \lambda\mathbf{I}| &= 0 \end{aligned}$$

and produces the characteristic equations for $\pm\omega\mathbf{B}\mathbf{A}^2$. Thus,

$$\lambda(\omega\hat{\mathbf{B}}\hat{\mathbf{A}}\hat{\mathbf{P}}\hat{\mathbf{A}}) = \pm\frac{\omega\lambda^2(\mathbf{A})}{1 + 2\omega\lambda(\mathbf{A})}. \tag{3.25}$$

Taking direct sums of the three terms (*i.e.,* taking one element from each set in (3.23), (3.24), and (3.25)) yields

$$\left\{\frac{\lambda(\mathbf{A})}{1 + 2\omega\lambda(\mathbf{A})}\right\} \oplus \left\{\frac{\omega\lambda^2(\mathbf{A})}{1 + 2\omega\lambda(\mathbf{A})}\right\} \oplus \left\{\frac{-\omega\lambda^2(\mathbf{A})}{1 + 2\omega\lambda(\mathbf{A})}, \frac{\omega\lambda^2(\mathbf{A})}{1 + 2\omega\lambda(\mathbf{A})}\right\}. \tag{3.26}$$

The sets of eigenvalues therefore are the sums of the set elements in (3.26) and the eigenvalues therefore are

$$\frac{\lambda(\mathbf{A})}{1 + 2\omega\lambda(\mathbf{A})} + \frac{\omega\lambda^2(\mathbf{A})}{1 + 2\omega\lambda(\mathbf{A})} - \frac{\omega\lambda^2(\mathbf{A})}{1 + 2\omega\lambda(\mathbf{A})} = \frac{\lambda(\mathbf{A})}{1 + 2\omega\lambda(\mathbf{A})}$$

and

$$\frac{\lambda(\mathbf{A})}{1 + 2\omega\lambda(\mathbf{A})} + \frac{\omega\lambda^2(\mathbf{A})}{1 + 2\omega\lambda(\mathbf{A})} + \frac{\omega\lambda^2(\mathbf{A})}{1 + 2\omega\lambda(\mathbf{A})} = \frac{[1 + 2\omega\lambda(\mathbf{A})]\lambda(\mathbf{A})}{1 + 2\omega\lambda(\mathbf{A})} = \lambda(\mathbf{A}).$$

Consequently, for each eigenvalue $\lambda_i(\mathbf{A})$ of $\mathbf{A}$, there are two eigenvalues of $\hat{\mathbf{C}}\hat{\mathbf{A}}$, namely, $\lambda_i(\mathbf{A})$ and $\frac{\lambda_i(\mathbf{A})}{1+2\omega\lambda_i(\mathbf{A})}$.

The statement $\left|\frac{\lambda_i(\mathbf{A})}{1+2\omega\lambda_i(\mathbf{A})}\right| < |\lambda_i(\mathbf{A})|$ follows from the fact that if $\mathbf{A}$ is positive (negative) definite, then $\lambda_i(\mathbf{A}) > 0$ ($\lambda_i(\mathbf{A}) < 0$). In this case, $\omega > 0$ ($\omega < 0$) and consequently the denominator $1 + 2\omega\lambda_i(\mathbf{A}) > 0$ and the result follows.

Statement ii: Using the same approach and notation as before with direct sums, the eigenvalues of $(1 - \omega)\hat{\mathbf{A}} + \omega\hat{\mathbf{A}}\hat{\mathbf{P}}$ can therefore be divided into two sets for calculating direct sums: the eigenvalues of $(1-\omega)\hat{\mathbf{A}}$ and the eigenvalues of $\omega\hat{\mathbf{A}}\hat{\mathbf{P}}$. From the proof in Statement i, the eigenvalues of $\hat{\mathbf{A}}$ are the same as the eigenvalues of $\mathbf{A}$. Hence, the eigenvalues of the first set in the direct summation is $(1 - \omega)\lambda(\mathbf{A})$.

The second set in the direct summation are the eigenvalues of $\omega\hat{\mathbf{A}}\hat{\mathbf{P}}$. Accounting for the permutation matrix in $\hat{\mathbf{A}}\hat{\mathbf{P}}$, the characteristic matrix is

$$\omega\hat{\mathbf{A}}\hat{\mathbf{P}} - \lambda\mathbf{I} = \begin{pmatrix} -\lambda\mathbf{I} & \omega\mathbf{A} \\ \omega\mathbf{A} & -\lambda\mathbf{I} \end{pmatrix}$$

and the characteristic equation is

$$\left|(-\lambda\mathbf{I})^2 - (\omega\mathbf{A})^2\right| = |(\lambda\mathbf{I} - \omega\mathbf{A})(\lambda\mathbf{I} + \omega\mathbf{A})| = 0$$

Multiplying each factor by -1 yields

$$\begin{aligned} |(\omega\mathbf{A} - \lambda\mathbf{I})(-\omega\mathbf{A} - \lambda\mathbf{I})| &= 0 \qquad\qquad \text{hence} \\ |\omega\mathbf{A} - \lambda\mathbf{I}||-\omega\mathbf{A} - \lambda\mathbf{I}| &= 0 \end{aligned}$$

and produces the characteristic equations for $\pm\omega\mathbf{A}$. Thus, $\lambda(\omega\hat{\mathbf{A}}\hat{\mathbf{P}}) = \pm\omega\lambda(\mathbf{A})$ and the components of the direct sums are

$$\{(1-\omega)\lambda(\mathbf{A})\} \oplus \{-\omega\lambda(\mathbf{A}), \omega\lambda(\mathbf{A})\}\,. \tag{3.27}$$

Therefore,

$$\begin{aligned} \lambda(\hat{\mathbf{A}} - \omega\hat{\mathbf{A}} + \omega\hat{\mathbf{A}}\hat{\mathbf{P}}) &= (1-w)\lambda(\mathbf{A}) \pm \omega\lambda(\mathbf{A}) \\ &= \{\lambda(\mathbf{A}), (1-2\omega)\lambda(\mathbf{A})\} \end{aligned} \tag{3.28}$$

It easily follows that if $\omega \in (-1/2, 1/2)$, then $|(1-2\omega)\lambda(\mathbf{A})| < |\lambda(\mathbf{A})|$.

Statement iii: To show that if $\mathbf{v}$ is an eigenvector of $\mathbf{A}$ then $(\mathbf{v},\mathbf{v})^T$ and $(\mathbf{v},-\mathbf{v})^T$ are eigenvectors of $\hat{\mathbf{C}}\hat{\mathbf{A}}$ and $\hat{\mathbf{A}}\hat{\mathbf{C}}'$, first consider the case for system $\hat{\mathbf{C}}\hat{\mathbf{A}}$ and let $\hat{\lambda}$ represent a generic eigenvalue of $\hat{\mathbf{C}}\hat{\mathbf{A}}$ such that

$$\hat{\mathbf{C}}\hat{\mathbf{A}}\hat{\mathbf{v}} = \hat{\lambda}\hat{\mathbf{v}} \tag{3.29}$$

for block vectors of the forms

$$\hat{\mathbf{v}} \in \left\{ \begin{bmatrix} \mathbf{v} \\ \mathbf{v} \end{bmatrix}, \begin{bmatrix} \mathbf{v} \\ -\mathbf{v} \end{bmatrix} \right\}.$$

Decomposing $\hat{\mathbf{C}}\hat{\mathbf{A}}$ in (3.29) and rearranging (see the form in (3.15)) we obtain

$$\hat{\mathbf{B}}(\hat{\mathbf{A}} + \omega\hat{\mathbf{A}}^2 + \omega\hat{\mathbf{A}}\hat{\mathbf{P}}\hat{\mathbf{A}})\hat{\mathbf{v}} = \hat{\lambda}\hat{\mathbf{v}}. \tag{3.30}$$

Multiplying (3.30) through by $\hat{\mathbf{B}}^{-1}$ then yields

$$\hat{\mathbf{A}}\hat{\mathbf{v}} + \omega\hat{\mathbf{A}}^2\hat{\mathbf{v}} + \omega\hat{\mathbf{A}}\hat{\mathbf{P}}\hat{\mathbf{A}}\hat{\mathbf{v}} = \hat{\lambda}\hat{\mathbf{B}}^{-1}\hat{\mathbf{v}}. \tag{3.31}$$

Considering the top rows of the matrices in (3.31) we obtain the following matrix equation:

$$\mathbf{A}\mathbf{v}_1 + \omega\mathbf{A}^2\mathbf{v}_1 + \omega\mathbf{A}^2\mathbf{v}_2 = \hat{\lambda}(\mathbf{I} + 2\omega\mathbf{A})\mathbf{v}_1 \tag{3.32}$$

where the vector elements of $\hat{\mathbf{v}}$ are identified with subscripts to indicate the operation of the permutation matrix $\hat{\mathbf{P}}$. The right-hand-side of (3.32) can be re-written in the following way:

$$\begin{aligned} \hat{\lambda}(\mathbf{I} + 2\omega\mathbf{A})\mathbf{v}_1 &= \hat{\lambda}\mathbf{v}_1 + 2\omega\hat{\lambda}\mathbf{A}\mathbf{v}_1 \\ &= \hat{\lambda}\mathbf{v}_1 + 2\omega\mathbf{A}\hat{\lambda}\mathbf{v}_1 && (3.33) \\ &= \mathbf{A}\mathbf{v}_1 + 2\omega\mathbf{A}^2\mathbf{v}_1 && (3.34) \end{aligned}$$

where we use the fact that if $\hat{\lambda} \in \lambda(\mathbf{A})$ in (3.33) then $\hat{\lambda}\mathbf{v}_1 = \mathbf{A}\mathbf{v}_1$ which gives (3.34). Combining this with the left-hand-side of (3.32) we obtain the following equation:

$$\mathbf{A}\mathbf{v}_1 + \omega\mathbf{A}^2\mathbf{v}_1 + \omega\mathbf{A}^2\mathbf{v}_2 = \mathbf{A}\mathbf{v}_1 + 2\omega\mathbf{A}^2\mathbf{v}_1. \tag{3.35}$$

Simplifying (3.35), we obtain,

$$\omega\mathbf{A}^2\mathbf{v}_2 = \omega\mathbf{A}^2\mathbf{v}_1$$

hence, $\mathbf{v}_1 = \mathbf{v}_2$, again, if the eigenvalues of $\hat{\mathbf{C}}\hat{\mathbf{A}}$ belong to the set $\lambda(\mathbf{A})$. If the eigenvalues are from the other set, *i.e.,* $\hat{\lambda} \in \frac{\lambda(\mathbf{A})}{1+2\omega\lambda(\mathbf{A})}$ then (3.33) leads to

$$\frac{\lambda(\mathbf{A})\mathbf{v}_1}{1+2\omega\lambda(\mathbf{A})} + \frac{2\omega\mathbf{A}\lambda(\mathbf{A})\mathbf{v}_1}{1+2\omega\lambda(\mathbf{A})} = \frac{\mathbf{A}\mathbf{v}_1}{1+2\omega\lambda(\mathbf{A})} + \frac{2\omega\mathbf{A}^2\mathbf{v}_1}{1+2\omega\lambda(\mathbf{A})} \tag{3.36}$$

where we use the fact that $\lambda(\mathbf{A})\mathbf{v}_1 = \mathbf{A}\mathbf{v}_1$. Combining (3.36) with the left-hand-side of (3.32) and simplifying, we obtain the equality

$$\omega\mathbf{A}^2\mathbf{v}_1 + \omega\mathbf{A}^2\mathbf{v}_2 = 0.$$

Hence, $\mathbf{v}_1 = -\mathbf{v}_2$.

Now consider the matrix $\hat{\mathbf{A}}\hat{\mathbf{C}}'$ in an input-based system. The proof of the eigenvectors proceeds in an identical fashion as above. Thus, for some eigenvector and eigenvalue, we have

$$(\hat{\mathbf{A}} - \omega\hat{\mathbf{A}} + \omega\hat{\mathbf{A}}\hat{\mathbf{P}})\hat{\mathbf{v}} = \hat{\lambda}\hat{\mathbf{v}}. \tag{3.37}$$

Taking the first row of matrices as before and using subscripts on the components of the eigenvector $\hat{\mathbf{v}}$ to indicate the effect of the permutation matrix we obtain

$$\mathbf{A}\mathbf{v}_1 - \omega\mathbf{A}\mathbf{v}_1 + \omega\mathbf{A}\mathbf{v}_2 = \hat{\lambda}\mathbf{v}_1. \tag{3.38}$$

If $\hat{\lambda} = \lambda(\mathbf{A})$, then (3.38) reduces to $\omega\mathbf{A}\mathbf{v}_1 = \omega\mathbf{A}\mathbf{v}_2$, hence $\mathbf{v}_1 = \mathbf{v}_2$. When $\hat{\lambda} = (1-2\omega)\lambda(\mathbf{A})$, then (3.38) reduces to $\omega\mathbf{A}\mathbf{v}_1 + \omega\mathbf{A}\mathbf{v}_2 = 0$, hence $\mathbf{v}_1 = -\mathbf{v}_2$. □

Corollary 1 to Theorem 3.2: *If in output-based systems, $\omega > 0$ (< 0) when $\mathbf{A}$ is positive (negative) definite, and if in input-based systems $\omega \in (-1/2, 1/2)$, then matrix $\mathbf{A}$ is diagonalizable if and only if matrices $\hat{\mathbf{C}}\hat{\mathbf{A}}$ and $\hat{\mathbf{A}}\hat{\mathbf{C}}'$ are diagonalizable.*

Proof. Assume matrix $\mathbf{A}$ is diagonalizable, hence has n linearly independent eigenvectors. Thus,

$$\sum_{i=1}^{n} \alpha_i \mathbf{v}_i = 0 \tag{3.39}$$

only when $\alpha_i = 0$ for all i.

To show that the matrices $\hat{\mathbf{C}}\hat{\mathbf{A}}$ and $\hat{\mathbf{A}}\hat{\mathbf{C}}'$ are diagonalizable, we must show that their eigenvectors are linearly independent. First note that from the theorem, the matrices $\hat{\mathbf{C}}\hat{\mathbf{A}}$ and $\hat{\mathbf{A}}\hat{\mathbf{C}}'$ must have $2n$ linearly independent eigenvectors for them to be diagonalizable. But from the theorem, every eigenvector $\mathbf{v}_i$ of $\mathbf{A}$ produces two eigenvectors of $\hat{\mathbf{C}}\hat{\mathbf{A}}$ and $\hat{\mathbf{A}}\hat{\mathbf{C}}'$, $[\mathbf{v}_i, \mathbf{v}_i]^T$ and $[\mathbf{v}_i, -\mathbf{v}_i]^T$ and therefore the matrices $\hat{\mathbf{C}}\hat{\mathbf{A}}$ and $\hat{\mathbf{A}}\hat{\mathbf{C}}'$ have $2n$ eigenvectors. Thus, we must show that

$$\sum_{i=1}^{n} \alpha_i' \begin{bmatrix} \mathbf{v}_i \\ \mathbf{v}_i \end{bmatrix} + \sum_{i=1}^{n} \beta_i' \begin{bmatrix} \mathbf{v}_i \\ -\mathbf{v}_i \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix} \tag{3.40}$$

only when $\alpha_i' = \beta_i' = 0$ for all i. Note that the bottom components sum to zero for any coefficients α_i' and β_i' when $\alpha_i' = \beta_i'$ for each i. Also, note that the top components in (3.40) sum to zero for any coefficients α_i and β_i when $\alpha_i = -\beta_i$ for each i. Together these two constraints, *i.e.,* if for all i, $\alpha_i = \beta_i$ and $\alpha_i = -\beta_i$ indicate that $\alpha_i = \beta_i = 0$ for all i as required.

Moreover, this is the only set of values satisfying (3.40). To see this, assume that there exists some non-zero values of α_i' and β_i' such that (3.40) is true. Then there is some vector that is a linear combination of the other $2n-1$ vectors. Say $\check{\mathbf{v}}_j$ is such a vector (similar reasoning can be applied to a vector $\tilde{\mathbf{v}}_j$). Then

$$\check{\mathbf{v}}_j = -\sum_{\substack{i=1 \\ i\neq j}}^{n} \frac{\alpha_i'}{\alpha_j'} \check{\mathbf{v}}_i - \sum_{i=1}^{n} \frac{\beta_i'}{\alpha_j'} \tilde{\mathbf{v}}_i \tag{3.41}$$

Rewriting (3.41) and summing just the top vector components and simplifying we get

$$\mathbf{v}_j = -\sum_{\substack{i=1 \\ i\neq j}}^{n} \left(\frac{\alpha_i' + \beta_i'}{\alpha_j' + \beta_j} \right) \mathbf{v}_i$$

contradicting the assumption of linear independence of the eigenvectors $\mathbf{v}_i$ of $\mathbf{A}$. Thus, the vectors $[\mathbf{v}_i, \mathbf{v}_i]^T$ and $[\mathbf{v}_i, -\mathbf{v}_i]^T$, $i = \{1, \dots, n\}$ are linearly independent and therefore the matrices $\hat{\mathbf{C}}\hat{\mathbf{A}}$ and $\hat{\mathbf{A}}\hat{\mathbf{C}}'$ are diagonalizable.

To prove the inverse, assume the eigenvectors of $\hat{\mathbf{C}}\hat{\mathbf{A}}$ and $\hat{\mathbf{A}}\hat{\mathbf{C}}'$ are diagonalizable, hence their eigenvectors are linearly independent. Then (3.40) is true only when $\alpha_i' = \beta_i' = 0$ for all $i = \{1, \dots, n\}$. It trivially follows then that (3.39) is also true, hence, the eigenvectors of $\mathbf{A}$ are linearly independent and therefore $\mathbf{A}$ is diagonalizable. □

Corollary 2 to Theorem 3.2: *Matrices $\hat{\mathbf{C}}\hat{\mathbf{A}}$ and $\hat{\mathbf{A}}\hat{\mathbf{C}}'$ have orthogonal eigenspaces if and only if matrix $\mathbf{A}$ has an orthogonal eigenspace.*

Proof. Assume matrix $\mathbf{A}$ has an orthogonal eigenspace. Then any two eigenvectors $\mathbf{v}_i$ and $\mathbf{v}_j$ of $\mathbf{A}$ are orthogonal hence $(\mathbf{v}_i^T \mathbf{v}_j) = 0$. From the theorem, each of these eigenvectors produces a pair of eigenvectors in the cross-coupled systems. Thus, for eigenvectors $\mathbf{v}_i$ and $\mathbf{v}_j$ of $\mathbf{A}$, both $\hat{\mathbf{C}}\hat{\mathbf{A}}$ and $\hat{\mathbf{A}}\hat{\mathbf{C}}'$ have eigenvectors

$$\check{\mathbf{v}}_i = \begin{bmatrix} \mathbf{v}_i \\ \mathbf{v}_i \end{bmatrix},\ \tilde{\mathbf{v}}_i = \begin{bmatrix} \mathbf{v}_i \\ -\mathbf{v}_i \end{bmatrix} \text{ and } \check{\mathbf{v}}_j = \begin{bmatrix} \mathbf{v}_j \\ \mathbf{v}_j \end{bmatrix},\ \tilde{\mathbf{v}}_j = \begin{bmatrix} \mathbf{v}_j \\ -\mathbf{v}_j \end{bmatrix},$$

respectively. Clearly, each vector pair $\check{\mathbf{v}}_i, \tilde{\mathbf{v}}_i$ are orthogonal, *i.e.,* for any i, $\check{\mathbf{v}}_i^T \tilde{\mathbf{v}}_i = 0$. It remains to show that vectors from two different pairs are also orthogonal. That is, $\check{\mathbf{v}}_i^T \check{\mathbf{v}}_j = 0$ and $\check{\mathbf{v}}_i^T \tilde{\mathbf{v}}_j = 0$. But

$$\check{\mathbf{v}}_i^T \check{\mathbf{v}}_j = [\mathbf{v}_i, \mathbf{v}_i]^T \begin{bmatrix} \mathbf{v}_j \\ \mathbf{v}_j \end{bmatrix} = (\mathbf{v}_i^T \mathbf{v}_j) + (\mathbf{v}_i^T \mathbf{v}_j) = (0 + 0) = 0,$$

and therefore are orthogonal. Also,

$$\begin{aligned}\check{\mathbf{v}}_i^T\tilde{\mathbf{v}}_j &= [\mathbf{v}_i, \mathbf{v}_i]^T \begin{bmatrix} \mathbf{v}_j \\ -\mathbf{v}_j \end{bmatrix} \\ &= (\mathbf{v}_i^T\mathbf{v}_j) + (\mathbf{v}_i^T(-\mathbf{v}_j)) \\ &= 0 + -0 = 0,\end{aligned}$$

hence are orthogonal. The other pairs of vectors $\tilde{\mathbf{v}}_i$ and $\hat{\mathbf{v}}_j$ are orthogonal by similar reasoning. Therefore, if all eigenvectors of $\mathbf{A}$ are orthogonal, all eigenvectors of $\hat{\mathbf{C}}\hat{\mathbf{A}}$ and $\hat{\mathbf{A}}\hat{\mathbf{C}}'$ are also orthogonal.

To prove the inverse, assume that for all eigenvectors of $\hat{\mathbf{C}}\hat{\mathbf{A}}$ and $\hat{\mathbf{A}}\hat{\mathbf{C}}'$ are orthogonal. Then for all $i, j = \{1, \ldots, n\}$ vectors $\check{\mathbf{v}}_i$ and $\check{\mathbf{v}}_j$ are orthogonal. Thus,

$$[\mathbf{v}_i, \mathbf{v}_i]^T \begin{bmatrix} \mathbf{v}_j \\ \mathbf{v}_j \end{bmatrix} = (\mathbf{v}_i^T\mathbf{v}_j) + (\mathbf{v}_i^T\mathbf{v}_j) = 2(\mathbf{v}_i^T\mathbf{v}_j) = 0$$

hence $(\mathbf{v}_i^T\mathbf{v}_j) = 0$ and $\mathbf{v}_i$ and $\mathbf{v}_j$ are orthogonal. Consequently, $\mathbf{A}$ has an orthogonal eigenspace. □

Corollary 3 to Theorem 3.2: *If in output-based systems, $\omega > 0$ (< 0) when $\mathbf{A}$ is positive (negative) definite, and if in input-based systems $\omega \in (-1/2, 1/2)$, then the spectral radius of matrix $\mathbf{A}$ equals the spectral radius of matrix $\hat{\mathbf{C}}\hat{\mathbf{A}}$ and the spectral radius of $\hat{\mathbf{A}}\hat{\mathbf{C}}'$,* i.e., *$\rho(\mathbf{A}) = \rho(\hat{\mathbf{C}}\hat{\mathbf{A}}) = \rho(\hat{\mathbf{A}}\hat{\mathbf{C}}')$.*

Proof. This is a direct implication of the fact that the largest eigenvalue of $\mathbf{A}$ is also an eigenvalue of $\hat{\mathbf{C}}\hat{\mathbf{A}}$ and $\hat{\mathbf{A}}\hat{\mathbf{C}}'$. Since for the values of ω indicated in the corollary statement, and from the theorem, all eigenvalues of $\hat{\mathbf{C}}\hat{\mathbf{A}}$ and $\hat{\mathbf{A}}\hat{\mathbf{C}}'$ are less than or equal to eigenvalues of $\mathbf{A}$, the equality of the spectral radii trivially follows. □

Example 1 illustrates Theorem 3.2 and it first corollary using a symmetric matrix.

Example 1. *Let*

$$\mathbf{A} = \begin{pmatrix} 6 & 2 \\ 2 & 3 \end{pmatrix},$$

a symmetric matrix, hence with an orthogonal eigenspace

$$\mathbf{v} \in \left\{ \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \end{pmatrix} \right\} \tag{3.42}$$

([3]) with eigenvalues $\{7, 2\}$ respectively. The matrices $\hat{\mathbf{C}}\hat{\mathbf{A}}$ and $\hat{\mathbf{A}}\hat{\mathbf{C}}'$ have an eigenspace of

$$\mathbf{v} \in \left\{ \begin{pmatrix} 2 \\ 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} -2 \\ -1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ -1 \\ 2 \end{pmatrix} \right\}$$

and the reader can verify their orthogonality. If $\omega = \frac{1}{2}$, then matrix $\hat{\mathbf{C}}\hat{\mathbf{A}}$ has eigenvalues $\{7, 2, \frac{7}{8}, \frac{2}{3}\}$. Note that the multiplier of the modified eigenvalues is $1/(1 +$

$2(.5)(7)) = 1/8$ *and* $1/(1+2(.5)(2)) = 1/3$ *for the eigenvalues* $7, 2$ *respectively. If* $\omega = \frac{1}{4}$ *then the multiplier for the modified eigenvalues in the corresponding input-based system is* $(1-2\omega) = 1/2$ *and matrix* $\hat{\mathbf{A}}\hat{\mathbf{C}}'$ *therefore has eigenvalues* $\{7, 2, \frac{7}{2}, 1\}$.

Before discussing the performance characteristics of cross-coupling and applications of Theorem 3.2, the following lemmas are needed.

LEMMA 3.3. *The vector* $\breve{\mathbf{x}}^* = [\mathbf{x}^*, \mathbf{x}^*]^T$ *is a fixed-point of the cross-coupled output-based system in (3.14),* i.e.,

$$\breve{\mathbf{x}}^* = (\mathbf{I} + \omega\hat{\mathbf{A}} - \omega\hat{\mathbf{A}}\hat{\mathbf{P}})^{-1}(\hat{\mathbf{A}}\breve{\mathbf{x}}^* + \breve{\mathbf{c}})$$

and in the cross-coupled input-based system in (3.18), i.e.,

$$\breve{\mathbf{x}}^* = \hat{\mathbf{A}}(\mathbf{I} - \omega\mathbf{I} + \omega\hat{\mathbf{P}})\breve{\mathbf{x}}^* + \breve{\mathbf{c}}$$

if and only if $\mathbf{x}^*$ *is a fixed-point of the system in (3.11),* i.e.,

$$\mathbf{x}^* = \mathbf{A}\mathbf{x}^* + \mathbf{c}.$$

Proof. First consider the output-based system. Assume $\mathbf{x}^*$ is a fixed-point solution for the system (3.11). From the definitions of $\hat{\mathbf{A}}$ and $\breve{\mathbf{c}}$, it is obvious that $\breve{\mathbf{x}}^* = \hat{\mathbf{A}}\breve{\mathbf{x}}^* + \breve{\mathbf{c}}$. Hence, we can substitute $\breve{\mathbf{x}}^*$ for $\hat{\mathbf{A}}\breve{\mathbf{x}}^* + \breve{\mathbf{c}}$ in the right-hand side of (3.14). Setting the output of (3.14) equal to $\hat{\mathbf{x}}'$, yields

$$\begin{aligned} \hat{\mathbf{x}}' &= (\mathbf{I} + \omega\hat{\mathbf{A}} - \omega\hat{\mathbf{A}}\hat{\mathbf{P}})^{-1}\breve{\mathbf{x}}^* \\ &= \hat{\mathbf{B}}(\mathbf{I} + \omega\hat{\mathbf{A}} + \omega\hat{\mathbf{A}}\hat{\mathbf{P}})\breve{\mathbf{x}}^*. \end{aligned} \tag{3.43}$$

Multiplying both sides of (3.43) by $\hat{\mathbf{B}}^{-1}$ (defined in (3.16) where its commutative properties are also described) we obtain

$$\begin{pmatrix} (\mathbf{I} + 2\omega\mathbf{A})\mathbf{x}_1' \\ (\mathbf{I} + 2\omega\mathbf{A})\mathbf{x}_2' \end{pmatrix} = (\mathbf{I} + \omega\hat{\mathbf{A}} + \omega\hat{\mathbf{A}}\hat{\mathbf{P}})\breve{\mathbf{x}}^*. \tag{3.44}$$

From the definition of $\breve{\mathbf{x}}^*$, $\hat{\mathbf{P}}\breve{\mathbf{x}}^* = \breve{\mathbf{x}}^*$ hence $\hat{\mathbf{A}}\hat{\mathbf{P}}\breve{\mathbf{x}}^* = \hat{\mathbf{A}}\breve{\mathbf{x}}^*$, and (3.44) becomes

$$\begin{pmatrix} (\mathbf{I} + 2\omega\mathbf{A})\mathbf{x}_1' \\ (\mathbf{I} + 2\omega\mathbf{A})\mathbf{x}_2' \end{pmatrix} = \begin{pmatrix} (\mathbf{I} + 2\omega\mathbf{A})\mathbf{x}_1^* \\ (\mathbf{I} + 2\omega\mathbf{A})\mathbf{x}_2^* \end{pmatrix} \tag{3.45}$$

and it easily follows that $\hat{\mathbf{x}}' = \breve{\mathbf{x}}^*$. Consequently, $\breve{\mathbf{x}}^*$ is a fixed point of system (3.14).

To show the inverse implication, let $\breve{\mathbf{x}}^*$ be a fixed point solution of (3.14). Therefore, multiplying both sides of (3.14) by the inverse of the cross-coupling matrix yields

$$(\mathbf{I} + \omega\hat{\mathbf{A}} - \omega\hat{\mathbf{A}}\hat{\mathbf{P}})\breve{\mathbf{x}}^* = \mathbf{A}\breve{\mathbf{x}}^* + \breve{\mathbf{c}}. \tag{3.46}$$

Again, noting that for $\breve{\mathbf{x}}^*$, $\hat{\mathbf{A}}\breve{\mathbf{x}}^* = \hat{\mathbf{A}}\hat{\mathbf{P}}\breve{\mathbf{x}}^*$, the left-hand side of (3.46) reduces to $\breve{\mathbf{x}}^*$. Therefore $\breve{\mathbf{x}}^* = \hat{\mathbf{A}}\breve{\mathbf{x}}^* + \breve{\mathbf{c}}$. Again, from the definitions of $\hat{\mathbf{A}}, \breve{\mathbf{x}}^*$ and $\breve{\mathbf{c}}$, it trivially follows that $\mathbf{x}^* = \mathbf{A}\mathbf{x}^* + \mathbf{c}$. Hence, $\mathbf{x}^*$ is a fixed point of (3.11).

Now consider the input-based system. Again assume that $\mathbf{x}^* = \mathbf{A}\mathbf{x}^* + \mathbf{c}$, a fixed point of (3.11). Then it follows that $\mathbf{x}^* = (1-\omega)\mathbf{A}\mathbf{x}^* + \omega\mathbf{A}\mathbf{x}^* + \mathbf{c}$ for all $0 \le \omega \le 1$. Thus, we can write two equations and put them in matrix vector form as follows:

$$\begin{aligned} \begin{bmatrix} \mathbf{x}^* \\ \mathbf{x}^* \end{bmatrix} &= \begin{bmatrix} (1-\omega)\mathbf{A}\mathbf{x}^* + \omega\mathbf{A}\mathbf{x}^* + \mathbf{c} \\ \omega\mathbf{A}\mathbf{x}^* + (1-\omega)\mathbf{A}\mathbf{x}^* + \mathbf{c} \end{bmatrix} \\ &= (\hat{\mathbf{A}} - \omega\hat{\mathbf{A}} + \omega\hat{\mathbf{A}}\hat{\mathbf{P}})\breve{\mathbf{x}}^* + \breve{\mathbf{c}} \end{aligned} \tag{3.47}$$

hence, by definition, $\breve{\mathbf{x}}^*$ is a fixed-point of (3.18).

To prove the inverse, assume that $\breve{\mathbf{x}}^*$ is a fixed-point of (3.47). Thus

$$\begin{aligned} \breve{\mathbf{x}}^* &= (\hat{\mathbf{A}} - \omega\hat{\mathbf{A}} + \omega\hat{\mathbf{A}}\hat{\mathbf{P}})\breve{\mathbf{x}}^* + \breve{\mathbf{c}} \\ &= \begin{bmatrix} (1-\omega)\mathbf{A}\mathbf{x}_1^* + \omega\mathbf{A}\mathbf{x}_2^* + \mathbf{c} \\ \omega\mathbf{A}\mathbf{x}_1^* + (1-\omega)\mathbf{A}\mathbf{x}_2^* + \mathbf{c} \end{bmatrix} \end{aligned}$$

and therefore each row in that matrix must be equal (here subscripts are used to distinguish vector components). Since these rows are equal it easily follows that $\mathbf{x}_1^* = \mathbf{x}_2^*$, hence $\mathbf{x}^* = \mathbf{A}\mathbf{x}^* + \mathbf{c}$ and is a fixed-point of system (3.11). ▯

LEMMA 3.4. *Let* $\mathbf{y}$ *be an* n *vector such that*

$$\mathbf{y} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i$$

for some linear combination of n *linearly independent vectors* $\mathbf{v}_i$. *Then* $\|\mathbf{y}\|_{\mathbf{V}} \equiv \sum_{i=1}^{n} |\alpha_i|$ *is a norm denominated as the norm with respect to the matrix* $\mathbf{V} = (\mathbf{v}_1|\mathbf{v}_2|\cdots|\mathbf{v}_n)$.

Proof. Recall that a norm has the following three properties:

1. Positivity: $\|\mathbf{y}\| \ge 0$ for all $\mathbf{y}$ and $\|\mathbf{y}\| = 0$ if and only if $\mathbf{y} = \mathbf{0}$.
2. Homogeneity: $\|c\mathbf{y}\| = |c|\|y\|$ for all scalars c and vectors $\mathbf{y}$.
3. The Triangle Inequality: $\|\mathbf{x}+\mathbf{y}\| \le \|\mathbf{x}\| + \|\mathbf{y}\|$ for all vectors $\mathbf{x}$ and $\mathbf{y}$.

The first property trivially follows from the linear independence of the vectors $\mathbf{v}_i$. Homogeneity can be proved by noting that

$$\begin{aligned} \|c\mathbf{y}\|_{\mathbf{V}} &= \left\| c\sum_{i=1}^{n} \alpha_i \mathbf{v}_i \right\|_{\mathbf{V}} = \left\| \sum_{i=1}^{n} c\alpha_i \mathbf{v}_i \right\|_{\mathbf{V}} \\ &= \sum_{i=1} |c\alpha_i| = \sum_{i=1} |c||\alpha_i| \\ &= |c| \sum_{i=1}^{n} |\alpha_i| = |c|\|\mathbf{y}\|_{\mathbf{V}} \end{aligned}$$

The Triangle Inequality is proved by showing that the norms with respect to matrix $\mathbf{V}$ for vectors $\mathbf{x} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i$ and $\mathbf{y} = \sum_{i=1}^{n} \beta_i \mathbf{v}_i$ are related by

$$\|\mathbf{x}+\mathbf{y}\|_{\mathbf{V}} = \left\| \sum_{i=1}^{n} \alpha_i \mathbf{v}_i + \sum_{i=1}^{n} \beta_i \mathbf{v}_i \right\|_{\mathbf{V}} = \left\| \sum_{i=1}^{n} (\alpha_i + \beta_i) \mathbf{v}_i \right\|_{\mathbf{V}}$$

$$= \sum_{i=1}^{n} |\alpha_i + \beta_i|$$
$$\leq \sum_{i=1}^{n} |\alpha_i| + \sum_{i=1}^{n} |\beta_i| = \|\mathbf{x}\|_{\mathbf{V}} + \|\mathbf{y}\|_{\mathbf{V}}$$

□

The value of Theorem 3.2 is that although the spectral radii of cross-coupled systems can never be less than that of matrix $\mathbf{A}$, it provides a means to *force* the reduction in the magnitude of eigenvalues governing iterates by using two inputs with opposite signs. This provides a general approach for determining the fixed point of a linear system: cross-coupling two equivalent systems and starting the iterations with two inputs of opposite signs. The following theorem shows that under certain simple conditions the cross-coupled system converges to the fixed points faster than the simple dynamical system.

The following definitions are needed:

Definition 3.5.

$\mathbf{V} = (\mathbf{v}_1|\mathbf{v}_2|\cdots|\mathbf{v}_n)$ *the matrix of the eigenvectors of* $\mathbf{A}$.

$\hat{\mathbf{V}} = (\check{\mathbf{v}}_1|\cdots|\check{\mathbf{v}}_n\|\tilde{\mathbf{v}}_1|\cdots|\tilde{\mathbf{v}}_n)$ *the matrix of the eigenvectors of a cross-coupled system.*

$\mathbf{M} \in \{\hat{\mathbf{C}}\hat{\mathbf{A}}, \hat{\mathbf{A}}\mathbf{C}'\}$ *The generic matrix* $\mathbf{M}$ *will be used to simplify the exposition. Its use indicates that either output-based or input-based cross-coupled systems and their respective matrices can be used within that particular context.*

$\mathbf{x}^{[0]}$ *the initial iterate in system (3.11).*

$\tilde{\mathbf{y}}^{[0]} = [\mathbf{y}^{[0]}, -\mathbf{y}^{[0]}]^T$ *the initial iterate in a cross-coupled system (either output or input-based).*

$\mathbf{x}^* = \sum_{i=1}^{n} \alpha_i' \mathbf{v}_i$ *for some set of* α_i'.

$\mathbf{e}_{\mathbf{x}}^{[0]} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i$ *for some set of* α_i.

$\check{\mathbf{e}}_{\mathbf{x}}^{[k]} \equiv \check{\mathbf{x}}^{[k]} - \check{\mathbf{x}}^*$ *the error vector in two independent (uncross coupled) systems (3.11).*

$\hat{\mathbf{e}}_{\mathbf{y}}^{[k]} \equiv \hat{\mathbf{y}}^{[k]} - \check{\mathbf{x}}^*$ *the error vector in a cross-coupled system.*

Theorem 3.6. *Let* $\mathbf{A}$ *be a diagonalizable matrix in system (3.11) and a matrix* $\mathbf{M}$ *be the corresponding matrix in its cross-coupled counterpart. If* $\mathbf{y}^{[0]} = \mathbf{x}^{[0]}$ *and* $\|\mathbf{x}^*\|_{\mathbf{V}} \ll \left\|\mathbf{e}_{\mathbf{x}}^{[0]}\right\|_{\mathbf{V}}$, i.e., *where for all* i

$$|\alpha_i'| < \left(\frac{\omega\lambda_i}{1+\omega\lambda_i}\right)|\alpha_i| \quad \text{and} \quad |\alpha_i'| < \left(\frac{\omega}{1-\omega}\right)|\alpha_i|$$

for output-based systems and input-based systems, respectively, then for all $k > 0$

$$\left\|\hat{\mathbf{e}}_{\mathbf{y}}^{[k]}\right\|_{\hat{\mathbf{V}}} < \left\|\mathbf{e}_{\mathbf{x}}^{[k]}\right\|_{\mathbf{V}} = \left\|\check{\mathbf{e}}_{\mathbf{x}}^{[k]}\right\|_{\hat{\mathbf{V}}}.$$

Proof. In this proof, a comparison is made between the convergence of the uncross-coupled system $\mathbf{x}^{[k+1]} = \mathbf{A}\mathbf{x}^{[k]} + \mathbf{c}$ and its cross-coupled counterparts. As the theorem statement indicates, this comparison is based on the norm defined in Lemma 3.4. This requires that the two systems being compared have the same dimension. The systems that we want to compare however have dimensions n and $2n$. Thus, it is necessary to convert the uncross-coupled system to an equivalent system with dimension $2n$. This is easily done by using two copies of the system in (3.11). Thus, the reader will note that the two independent dynamical systems in $\breve{\mathbf{x}}^{[k+1]} = \hat{\mathbf{A}}\breve{\mathbf{x}}^{[k]} + \breve{\mathbf{c}}$ evolves in an identical way as does the original dynamical system in the sense that the iterates of one imply the iterates of the other *i.e.,* the vector components in $\breve{\mathbf{x}}^{[k]}$ are equal to $\mathbf{x}^{[k]}$, so long as the value of $\mathbf{x}^{[0]}$ is the same in both systems which the theorem statement requires.

This approach to making comparisons is valid because $\|\breve{\mathbf{e}}_{\mathbf{x}}^{[k]}\|_{\hat{\mathbf{V}}} = \|\mathbf{e}_{\mathbf{x}}^{[k]}\|_{\mathbf{V}}$. To see this, define the iterated error vector for the linear system in (3.11) by $\mathbf{e}_{\mathbf{x}}^{[k+1]} = \mathbf{A}\mathbf{e}_{\mathbf{x}}^{[k]}$ hence

$$\mathbf{e}_{\mathbf{x}}^{[k]} = \mathbf{A}^k\mathbf{e}_{\mathbf{x}}^{[0]} = \mathbf{A}^k\sum_{i=1}^{n}\alpha_i\mathbf{v}_i = \sum_{i=1}^{n}\alpha_i\mathbf{A}^k\mathbf{v}_i = \sum_{i=1}^{n}\alpha_i\lambda_i^k\mathbf{v}_i. \tag{3.48}$$

Applying Lemma 3.4, then $\left\|\mathbf{e}_{\mathbf{x}}^{[k]}\right\|_{\mathbf{V}} = \sum_{i=1}^{n}\left|\alpha_i\lambda_i^k\right|$. By analogous arguments,

$$\breve{\mathbf{e}}_{\mathbf{x}}^{[k]} = \sum_{i=1}^{n}\alpha_i\lambda_i^k\breve{\mathbf{v}}_i \tag{3.49}$$

where the α_i are the same in (3.48) and (3.49) and therefore

$$\left\|\breve{\mathbf{e}}_{\mathbf{x}}^{[k]}\right\|_{\hat{\mathbf{V}}} = \sum_{i=1}^{n}\left|\alpha_i\lambda_i^k\right| = \left\|\mathbf{e}_{\mathbf{x}}^{[k]}\right\|_{\mathbf{V}}. \tag{3.50}$$

Before proceeding further with the main elements of the proof, it is helpful to develop the following expressions for the norm of the k^{th} iterate in cross-coupled systems. Despite the differences between the definitions of output-based and input-based systems (the matrix $\hat{\mathbf{C}}\hat{\mathbf{A}}$ is distributed whereas the matrix $\hat{\mathbf{A}}\hat{\mathbf{C}}'$ is not) the iterated error vector for either cross-coupled system can be expressed by

$$\hat{\mathbf{e}}_{\mathbf{y}}^{[k]} = \mathbf{M}^k\hat{\mathbf{e}}_{\mathbf{y}}^{[0]}. \tag{3.51}$$

It is useful to decompose the vector $\hat{\mathbf{e}}_{\mathbf{y}}^{[0]}$. Thus,

$$\hat{\mathbf{e}}_{\mathbf{y}}^{[0]} = \tilde{\mathbf{y}}^{[0]} - \breve{\mathbf{x}}^* = \begin{bmatrix}\mathbf{y}^{[0]}\\ -\mathbf{y}^{[0]}\end{bmatrix} - \begin{bmatrix}\mathbf{x}^*\\ \mathbf{x}^*\end{bmatrix}. \tag{3.52}$$

Letting $\mathbf{e}_{\mathbf{y}}^{[0]} = \mathbf{y}^{[0]} - \mathbf{x}^*$ then (3.52) becomes

$$\hat{\mathbf{e}}_{\mathbf{y}}^{[0]} = \begin{bmatrix}\mathbf{e}_{\mathbf{y}}^{[0]}\\ -\mathbf{e}_{\mathbf{y}}^{[0]}\end{bmatrix} - \begin{bmatrix}\mathbf{0}\\ 2\mathbf{x}^*\end{bmatrix} = \tilde{\mathbf{e}}_{\mathbf{y}}^{[0]} - 2\begin{bmatrix}\mathbf{0}\\ \mathbf{x}^*\end{bmatrix}. \tag{3.53}$$

Consequently, the k^{th} error vector in cross-coupled systems

$$\hat{\mathbf{e}}_{\mathbf{y}}^{[k]} = \mathbf{M}^k\hat{\mathbf{e}}_{\mathbf{y}}^{[0]} = \mathbf{M}^k\tilde{\mathbf{e}}_{\mathbf{y}}^{[0]} - 2\mathbf{M}^k\begin{bmatrix}\mathbf{0}\\ \mathbf{x}^*\end{bmatrix}. \tag{3.54}$$

Now, since $\mathbf{A}$ is diagonalizable, then by the first corollary of Theorem 3.2, both matrices in set $\mathbf{M}$ are diagonalizable. Thus, any block vector can be written as a linear combination of the eigenvectors of $\mathbf{M}$. Thus, in (3.54)

$$\tilde{\mathbf{e}}_{\mathbf{y}}^{[0]} = \begin{bmatrix} \mathbf{e}_{\mathbf{y}}^{[0]} \\ -\mathbf{e}_{\mathbf{y}}^{[0]} \end{bmatrix} = \sum_{i=1}^{n} \alpha_i \begin{bmatrix} \mathbf{v}_i \\ -\mathbf{v}_i \end{bmatrix} \tag{3.55}$$

where since $\mathbf{y}^{[0]} = \mathbf{x}^{[0]}$ the α_i in (3.55) are the same as those in (3.50). The first term on the right in (3.54) therefore is

$$\mathbf{M}^k \tilde{\mathbf{e}}_{\mathbf{y}}^{[0]} = \sum_{i=1}^{n} \alpha_i \mathbf{M}^k \begin{bmatrix} \mathbf{v}_i \\ -\mathbf{v}_i \end{bmatrix} = \sum_{i=1}^{n} \alpha_i \hat{\lambda}_i^k \begin{bmatrix} \mathbf{v}_i \\ -\mathbf{v}_i \end{bmatrix}$$

where either $\hat{\lambda}_i = \frac{\lambda_i}{1+2\omega\lambda_i}$ or $\hat{\lambda}_i = (1-2\omega)\lambda_i$ from Theorem 3.2. In the second term of (3.54),

$$\begin{bmatrix} \mathbf{0} \\ \mathbf{x}^* \end{bmatrix} = \sum_{i=1}^{n} \alpha_i' \begin{bmatrix} \mathbf{0} \\ \mathbf{v}_i \end{bmatrix} = \sum_{i=1}^{n} \frac{\alpha_i'}{2} \begin{bmatrix} \mathbf{v}_i \\ \mathbf{v}_i \end{bmatrix} - \sum_{i=1}^{n} \frac{\alpha_i'}{2} \begin{bmatrix} \mathbf{v}_i \\ -\mathbf{v}_i \end{bmatrix} \tag{3.56}$$

for some α_i'. The second term on the right in (3.54) therefore becomes

$$\begin{aligned} 2\mathbf{M}^k \begin{bmatrix} \mathbf{0} \\ \mathbf{x}^* \end{bmatrix} &= 2\mathbf{M}^k \sum_{i=1}^{n} \frac{\alpha_i'}{2} \check{\mathbf{v}}_i - 2\mathbf{M}^k \sum_{i=1}^{n} \frac{\alpha_i'}{2} \tilde{\mathbf{v}}_i \\ &= \sum_{i=1}^{n} \alpha_i' \lambda_i^k \check{\mathbf{v}}_i - \sum_{i=1}^{n} \alpha_i' \hat{\lambda}_i^k \tilde{\mathbf{v}}_i. \end{aligned}$$

Thus, from (3.54), the norm

$$\begin{aligned} \left\| \hat{\mathbf{e}}_{\mathbf{y}}^{[k]} \right\|_{\hat{\mathbf{V}}} &= \left\| \sum_{i=1}^{n} \alpha_i \hat{\lambda}_i^k \tilde{\mathbf{v}}_i - \sum_{i=1}^{n} \alpha_i' \lambda_i^k \check{\mathbf{v}}_i + \sum_{i=1}^{n} \alpha_i' \hat{\lambda}_i^k \tilde{\mathbf{v}}_i \right\|_{\hat{\mathbf{V}}} \\ &= \left\| \sum_{i=1}^{n} (\alpha_i + \alpha_i') \hat{\lambda}_i^k \tilde{\mathbf{v}}_i - \sum_{i=1}^{n} \alpha_i' \lambda_i^k \check{\mathbf{v}}_i \right\|_{\hat{\mathbf{V}}} \\ &= \sum_{i=1}^{n} \left| (\alpha_i + \alpha_i') \hat{\lambda}_i^k \right| + \sum_{i=1}^{n} \left| \alpha_i' \lambda_i^k \right|. \end{aligned} \tag{3.57}$$

With these expressions for the norms taken care of, the following proof shows that if $\|\mathbf{x}^*\| \ll \|\mathbf{e}_{\mathbf{x}}^{[0]}\|$ by ensuring that each $|\alpha_i'| \ll |\alpha_i'|$ then the norm of the errors of cross-coupled systems are always less than that for single systems. Thus, for output-based systems (this is to save space since the proof for input-based systems follows analogously) and for all i,

$$\begin{aligned} |\alpha_i'| < \left(\frac{\omega\lambda_i}{1+\omega\lambda_i} \right) |\alpha_i| &= \left(\frac{2\omega\lambda_i}{2+2\omega\lambda_i} \right) |\alpha_i| \\ &= \left(\frac{1+2\omega\lambda_i - 1}{1+2\omega\lambda_i + 1} \right) |\alpha_i| \\ &\leq \left(\frac{1 - \frac{1}{(1+2\omega\lambda_i)^k}}{1 + \frac{1}{(1+2\omega\lambda_i)^k}} \right) |\alpha_i| \text{ for } k \geq 1 \text{, but} \end{aligned} \tag{3.58}$$

$$(3.59) \qquad \frac{1}{(1+2\omega\lambda_i)^k} = \frac{|\lambda_i^k|/(1+2\omega\lambda_i)^k}{|\lambda_i^k|} = \frac{|\hat{\lambda}_i^k|}{|\lambda_i^k|}.$$

Substituting (3.59) into (3.58) and rearranging we get

$$|\alpha_i'| < \left(\frac{|\lambda_i^k| - \left|\hat{\lambda}_i^k\right|}{|\lambda_i^k| + \left|\hat{\lambda}_i^k\right|} \right) |\alpha_i|$$

leading to

$$(3.60) \qquad \left|\alpha_i \hat{\lambda}_i^k\right| + \left|\alpha_i' \hat{\lambda}_i^k\right| + \left|\alpha_i' \lambda_i^k\right| < \left|\alpha_i \lambda_i^k\right|.$$

Combining terms on the left of (3.60)

$$\left|(\alpha_i + \alpha_i')\hat{\lambda}_i^k\right| + \left|\alpha_i' \lambda_i^k\right| \leq \left|\alpha_i \hat{\lambda}_i^k\right| + \left|\alpha_i' \hat{\lambda}_i^k\right| + \left|\alpha_i' \lambda_i^k\right| < \left|\alpha_i \lambda_i^k\right|.$$

Summing terms and for all $k > 0$ yields the desired result

$$\sum_{i=1}^{n} \left|(\alpha_i + \alpha_i')\hat{\lambda}_i^k\right| + \sum_{i=1}^{n} \left|\alpha_i' \lambda_i^k\right| = \left\|\tilde{\mathbf{e}}_{\mathbf{y}}^{[k]}\right\|_{\hat{\mathbf{V}}} < \sum_{i=1}^{n} \left|\alpha_i \lambda_i^k\right| = \left\|\mathbf{e}_{\mathbf{x}}^{[k]}\right\|_{\mathbf{V}}.$$

□

There are ways to obviate this condition and thereby ensure that *only* modified eigenvalues govern the iterates. This is the subject of a follow-on paper. The following example illustrates the effects of the condition $\|\mathbf{x}^*\| \ll \|\mathbf{e}_{\mathbf{x}}^{[0]}\|$.

EXAMPLE 2. *Consider the simple fixed point equation $x^{[k+1]} = \frac{1}{2}x^{[k]} + 1$ which converges to the fixed point $x^* = 2$. In Table 3.1, the output-based cross-coupled system with $\omega = 10$ converges more slowly than the single system. The effect of the initial inputs of opposite sign in the cross-coupled system when one is initially close to the fixed point slows down the convergence. One system pulls the other system too far away from a good estimate of the fixed point.*

In Table 3.2, the initial inputs are relatively large in magnitude compared to the magnitude of the fixed point. The beneficial effects of the inputs with opposite signs in getting the modified eigenvalues to operate on the iterates is evident. In higher dimensions, the cross-coupled system may offer real advantages in computing fixed points.

4. Scaling: Cross-Coupling Cross-Coupled Systems. In both output-based and input-based cross-coupling, it is possible to render the cross-coupled system into another feedforward system or linear iterative map of higher dimension for certain classes of matrices $\mathbf{A}$, hereafter referred to as the *root matrix*. This suggests the possibility of cross-coupling these cross-coupled systems and producing yet another, higher order, cross-coupled system. Indeed, this can be carried out *ad infinitum*. The question arises as what are its eigenvalues and how many of them are produced? In this section, some patterns of multiply cross-coupled systems and related scaling phenomena are explored.

It is useful to clarify how many times a system is cross coupled. The following convention will be used: the designation of an *order number m* will indicate how

Iteration	Single System	Cross-Coupled System	
0	2.25	2.25	-2.25
1	2.125	1.102272727	0.897727273
2	2.0625	1.50464876	1.49535124
3	2.03125	1.750211307	1.749788693
4	2.015625	1.875009605	1.874990395
5	2.0078125	1.937500437	1.937499563
6	2.00390625	1.96875002	1.96874998
7	2.001953125	1.984375001	1.984374999
8	2.000976563	1.9921875	1.9921875
9	2.000488281	1.99609375	1.99609375
10	2.000244141	1.998046875	1.998046875

Table 3.1
Comparison when $\|\mathbf{x}^*\|$ *is large relative to* $\left\|\mathbf{e}_{\mathbf{x}}^{[0]}\right\|$.

Iteration	Single System	Cross-Coupled System	
0	100	100	-100
1	51	5.545454545	-3.545454545
2	26.5	1.70661157	1.29338843
3	14.25	1.759391435	1.740608565
4	8.125	1.875426883	1.874573117
5	5.0625	1.937519404	1.937480596
6	3.53125	1.968750882	1.968749118
7	2.765625	1.98437504	1.98437496
8	2.3828125	1.992187502	1.992187498
9	2.19140625	1.99609375	1.99609375
10	2.095703125	1.998046875	1.998046875

Table 3.2
Comparison when $\|\mathbf{x}^*\|$ *is small relative to* $\left\|\mathbf{e}_{\mathbf{x}}^{[0]}\right\|$.

many times a system is cross-coupled. Thus, an order number of $m = 0$ refers to the original uncross-coupled system with the root matrix $\mathbf{A}$. An order number of $m = 1$ means a system of order 0 is cross-coupled and its transformation matrix is $\hat{\mathbf{C}}\hat{\mathbf{A}}$ or $\hat{\mathbf{A}}\hat{\mathbf{C}}'$ depending on whether it is an output-based or input-based system. So far, we have considered only systems of order $m = 1$.

The eigenvalues and eigenvectors of such multiply cross-coupled systems can be determined by recursive application of Theorem 3.2. The following theorem indicates a general form for these eigenvalues and their multiplicity.

Theorem 4.1. *Every eigenvalue in an m-order cross-coupled system has the form*

$$\frac{\lambda_i(\mathbf{A})}{1 + 2\omega k \lambda_i(\mathbf{A})} \quad \text{or} \quad (1-2\omega)^k \lambda_i(\mathbf{A}) \quad \text{for} \quad k = \{0, \ldots, m\}$$

with a multiplicity $r_{m,k} = \binom{m}{k}$ *for* $k = \{0, \ldots, m\}$ *for output-based and input-based systems, respectively.*

Proof. This is proved by induction using recursive application of Theorem 3.2. First note that for $k = 0$, the general form reduces to $\lambda_i(\mathbf{A})$ and that for any m, $\binom{m}{0} = 1$ as expected. Also, the general form reduces to the modified eigenvalue with $k = 1$ and $m = 1$. If $m = 1$, the number of eigenvalues spawned by $\lambda_i(\mathbf{A})$ are $\binom{1}{0} = 1$ for $\lambda_i(\mathbf{A})$ and $\binom{1}{1} = 1$ for the reduced eigenvalue. Further, there are no eigenvalues of the general form with $k > 1$ when $m = 1$. Thus, the general form holds for $m = 1$ and $k = \{0, 1\}$.

To show the general form holds for arbitrary m, we recursively applying Theorem 3.2 which states that any eigenvalue defined for some positive integer k produces two eigenvalues in the next order value when m is incremented: itself where the value of k stays the same, and the modified eigenvalue of the form

$$\frac{\frac{\lambda_i(\mathbf{A})}{1+2k\omega\lambda_i(\mathbf{A})}}{1 + 2\omega\left(\frac{\lambda_i(\mathbf{A})}{1+2k\omega\lambda_i(\mathbf{A})}\right)} = \frac{\frac{\lambda_i(\mathbf{A})}{1+2k\omega\lambda_i(\mathbf{A})}}{\frac{1+2k\omega\lambda_i(\mathbf{A})+2\omega\lambda_i(\mathbf{A})}{1+2k\omega\lambda_i(\mathbf{A})}} = \frac{\lambda_i(\mathbf{A})}{1 + 2(k+1)\omega\lambda_i(\mathbf{A})}$$

where the value of k is incremented. By similar reasoning and recursion, this is also true for input-based systems. That is, if the eigenvalue in question is $(1 - 2\omega)^k \lambda_i(\mathbf{A})$, then its modified child in the next order is

$$(1 - 2\omega)\left[(1 - 2\omega)^k \lambda_i(\mathbf{A})\right] = (1 - 2\omega)^{k+1}\lambda_i(\mathbf{A})$$

From this, the total number of eigenvalues in an m-order system associated with the index $k = \{0, \ldots, m\}$ is the sum of the number of those eigenvalues in order $m - 1$ with index $k - 1$ (which produces the modified eigenvalues with index k) and the number of those eigenvalues in order $m - 1$ with index k (which produce copies of themselves, *i.e.,* with index k). Applying the general form for the number of these eigenvalues in order $m-1$, the total number of eigenvalues in an m order system with index k is therefore given by the well-known recurrence relation

$$r_{m,k} = \binom{m-1}{k-1} + \binom{m-1}{k} = \binom{m}{k}$$

Thus, the general form holds for the next order m. □

Figure 4.1 illustrates this by showing how the recursive application of Theorem 3.2 results in a specific spectrum of eigenvalues for an output-based system. The number of times a system is cross-coupled is indicated by the order number at the beginning of each row. The eigenvalues in that row therefore correspond to the spectrum of eigenvalues for a cross-coupled system of that order. The arrows point to eigenvalues produced by cross-coupling and application of the theorem. The form of the multiplicities indicates that these multiplicities have the pattern associated with *Pascal's Triangle*. The number of arrows pointing to a particular eigenvalue indicates its multiplicity. Figure 4.2 illustrates the effect on eigenvectors of higher-order cross-coupling.

Observe that with every cross-coupling, the number of distinct eigenvalues associated with a single kernel increases by one: from 2 in a first-order system to 3 in the second-order system, but the *number* of eigenvalues has doubled as expected since the dimensions of the matrices double with every cross-coupling.

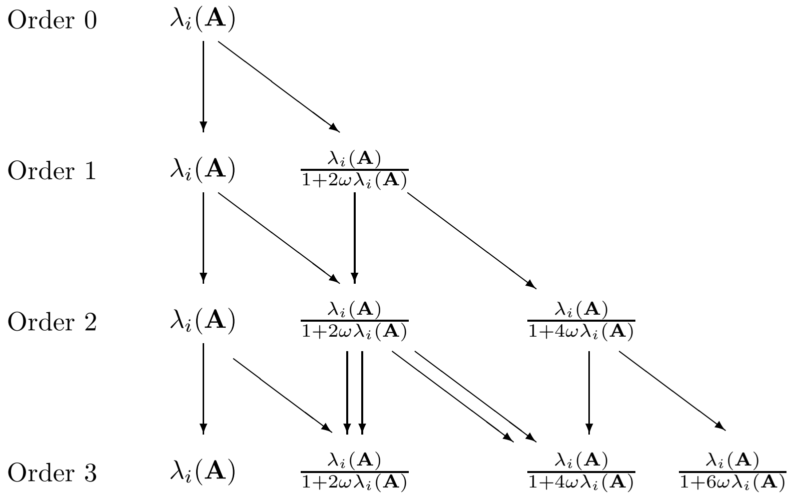

FIG. 4.1. *Pascal's Triangle Showing the Multiplicities of Eigenvalues*

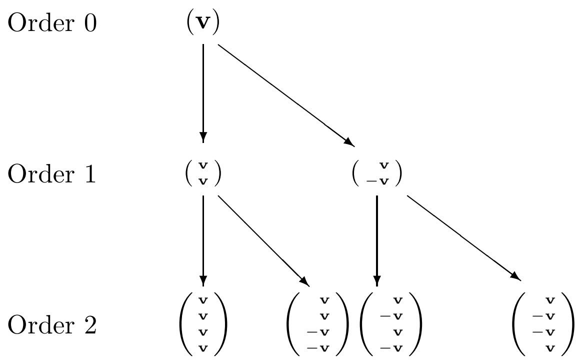

FIG. 4.2. *Pascal's Triangle Showing the Corresponding Eigenvectors for Orders 0-2*

Finally, it is worth showing the structure of a second order cross-coupled input-based system (the structure of the corresponding output system is quite cumbersome and less elegant). Using the matrix in (3.19) as the kernel, then the structure of the second-order cross-coupled input system is obtained by recursively substituting (3.19) for the matrix $\mathbf{A}$. This yields

$$\begin{pmatrix} (1-\omega)^2\mathbf{A} & (1-\omega)\omega\mathbf{A} & \omega(1-\omega)\mathbf{A} & \omega^2\mathbf{A} \\ (1-\omega)\omega\mathbf{A} & (1-\omega)^2\mathbf{A} & \omega^2\mathbf{A} & \omega(1-\omega)\mathbf{A} \\ \omega(1-\omega)\mathbf{A} & \omega^2\mathbf{A} & (1-\omega)^2\mathbf{A} & (1-\omega)\omega\mathbf{A} \\ \omega^2\mathbf{A} & \omega(1-\omega)\mathbf{A} & (1-\omega)\omega\mathbf{A} & (1-\omega)^2\mathbf{A} \end{pmatrix}$$

and upon close inspection, the sums of the rows and columns add up to the corresponding rows and columns of $\mathbf{A}$ and a pattern in the entries is evident. The significance of this is if $\mathbf{A}$ is a stochastic or doubly stochastic matrix, then all higher order matrices are also stochastic.

5. Applications and Future Research. One immediate if not obvious application of the type of cross-coupling described here is in the application to iterative methods for the solution of linear systems of equations. One popular method is the Gauss-Seidel (GS) iterative technique. Cross-coupling can be used in conjunction with GS and broaden its applicability, *i.e.,* remove some of the limitations of the GS method.

GS is an iterative method in which, given a linear system $\mathbf{Ax} = \mathbf{b}$, the value $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ is to be determined. This is done by creating an equivalent dynamical system that converges to the solution. By adding a vector $\mathbf{Bx}$ for some matrix $\mathbf{B}$ to both sides of $\mathbf{Ax} = \mathbf{b}$ and after some simple algebra, the system becomes $\mathbf{x} = (\mathbf{I} - \mathbf{B}^{-1}\mathbf{A})\mathbf{x} + \mathbf{B}^{-1}\mathbf{b}$. This results in the vector $\mathbf{x}$ on both sides, hence, a dynamical system. For suitable matrices $\mathbf{B}$, this system converges to the fixed-point solution $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$.

The problem with this general approach is in finding a suitable matrix $\mathbf{B}$. GS in particular uses successive updates for each vector element. This is equivalent to a matrix $\mathbf{B}$ defined as the lower triangular part of $\mathbf{A}$. In this case, the matrix $(\mathbf{I} - \mathbf{B}^{-1}\mathbf{A})$ is *diagonally dominant* with spectral radius $\rho((\mathbf{I} - \mathbf{B}^{-1}\mathbf{A})) < 1$ guaranteeing convergence. Many other *relaxation* methods constitute variations on the GS method [6].

This type of dynamical system can easily be cross-coupled. The kernel matrix is $(\mathbf{I} - \mathbf{B}^{-1}\mathbf{A})$ and we need not have any particular requirement on matrix $\mathbf{B}$ other than non-singularity. An initial iterate $\mathbf{x}^{[0]}$ and an *auxiliary iterate* $-\mathbf{x}^{[0]}$ provide the initial inputs to the cross-coupled system. Choosing appropriate values of ω and utilizing Theorem 3.6, a cross-coupled GS would converge faster and have a larger domain of applicability than simple GS (obviously, some such matrices $\mathbf{B}$ will present more or less round-off and truncation errors and the like, but we can ignore that issue for the time being). Thus, for some values of ω, a system where $\lambda(\mathbf{A}) > 1$ can still produce a cross-coupled system where $\frac{\lambda(\mathbf{A})}{1+2\omega\lambda(\mathbf{A})} < 1$ for output-based systems and $(1 - 2\omega)\lambda(\mathbf{A}) < 1$ for input-based systems, hence, converge to some solution notwithstanding the fact that $\lambda(\mathbf{A}) > 1$.

Indeed, the general concept of cross-coupling systems as described here, where each system of a pair provides an approximation of a reference signal, *i.e.,* some limiting value, could be applied to many dynamical systems, linear as well as non-linear. Future research might show some interesting forms for eigenvalues in non-linear dynamical systems such as the logistic function, a quadratic form similar to the linear systems considered here.

The performance of neural networks could be improved leading to better methods for pattern recognition, etc. In fact, the scaling property provides the intriguing possibility of a system that *adaptively modulates the order of cross-coupling*. Such modulation would then modulate the convergence rates of the dynamical system. It is indeed conceivable that this may provide a means of implementing an *attentional system* whereby more or less processing power is applied to a variety of simultaneous pattern recognition problems.

Probabilistic methods may also benefit. Recall that in input-based systems, the column and row sums are preserved. This suggests that stochastic matrices could be augmented by cross-coupling and still yield a stochastic matrix. It would be interesting therefore to see how this approach could be applied to Markov Decision Processes, a type of finite improvement algorithm.

Finally, it would be very interesting to examine under what system constraints cross-coupling is an optimal control strategy. Answers to this question could provide enormous value to the theory of autonomous systems and cognitive systems. If nature is indeed efficient, perhaps this efficiency can be described in terms of cross-coupling and lead to further development and understanding of cognitive systems.

6. Discussion and Conclusion. This article has provided some basic mathematical results on the dynamics of linear cross-coupled systems. These systems were simple linear systems designed to converge to a fixed point solution. These linear, cross-coupled systems provide a very simple model motivated by notions of autonomous control. This perspective allowed us to consider using a parallel system to provide an approximation of a reference signal that was used to approximate the error between an iterate and the fixed point.

In the context of the human brain, many outstanding issues regarding the notion of cross-coupling will no doubt exist for a long time to come. Issues regarding the size, structure and possibly different types of cross-coupling need to be explored. It may be that each hemisphere requires a structure similar in size and complexity to effect a sufficient level of feedback (see [8] where a distinction is made between *neuron-centric* feedback and *network-centric* feedback). Perhaps nature has provided a means for *optimal control* given the constraints imposed by biology, memory, etc. Perhaps the two systems provide sufficient information to each other to effect the control necessary in autonomous systems. Control via cross-coupling at the system level may therefore be the most efficient and robust way to accomplish this. Indeed, there may be a fractal-like hierarchy of cross-coupling which may point the way toward solving how brains can help steer a person through a randomly changing environment.

Two types of cross-coupled systems were analyzed: one based on using output values are a reference approximation; the other used input values (or the output values of the prior iteration). Analysis of both cross-coupled systems showed that when inputs to each system are the same, the outputs converge at the same rate as the original uncross-coupled system. When the inputs differed by a sign change, however, the system converges at a faster rate. These conclusions were based on a spectral analysis of the eigenvalues of a cross-coupled system. Cross-coupling doubles the size of the system matrix, doubles the number of eigenvalues and increases by one the number of distinct eigenvalues. In a cross-coupled system, each eigenvalue of the original system maps into two eigenvalues of the cross-coupled system—a copy of the original eigenvalue and a modification of the original eigenvalue. For a certain range of values of a parameter ω, these modified eigenvalues are guaranteed to have a smaller magnitude than the original eigenvalues. A theorem was presented that showed how these modified eigenvalues can be forced to control the dynamics of the system.

A scaling and self-similar structure of cross-coupled systems was also examined. Analysis showed that for higher order cross-coupling, certain properties of the transformation matrix are preserved. A general form for the eigenvalues was also described and the multiplicities of the eigenvalues was illustrated using Pascal's Triangle.

These results show how cross-coupling could be applied to many different types of linear systems. The analysis showed how it is possible to force the reduction of the eigenvalues that govern the iterative map by using two inputs differing by a sign change. Applications to numerical methods such as iterative methods for the solution of linear systems are therefore worth examining.

The potential applications of cross-coupling was also described and may be richer than those associated with linear iterative maps. The simple linear systems that were analyzed here provide clues as to how this approach might be applied to other problem domains such as optimal control, cognitive systems, neural networks, Markov Decision Processes, and stochastic control methods, to name but a few. Indeed, even for the simple systems described here, more work needs to be done to articulate all of their interesting properties. This article has only articulated some of them.

Finally, the motivation for this article may lead to new ways to view self-referencing systems. The exploitation of the paradigm of two identical (or perhaps similar) systems providing each other with reference signals may lead to complex systems capable of higher degrees of autonomy. Like the two hemispheres of the brain, ways of linking processing systems together may lead to better robotics, control and information systems. Afterall, as everyone knows, two heads are better than one.

REFERENCES

[1] D. J. Amit, *Modeling Brain Function: The World of Attractor Neural Networks*, The Cambridge University Press, New York, NY., 1989.

[2] Y. Bar-Yam, *Dynamics of Complex Systems*, Perseus Books, Reading, MA., 1997.

[3] R. Bellman, *Introduction to Matrix Analysis*, McGraw-Hill Book Company, New York, NY, 2 ed., 1970.

[4] F. T. S. Committee, ed., *Fed-Std-1037C 1996, Telecommunications: Glossary of Telecommunications Terms*, U.S. General Services Administration, 1996.

[5] F. H. C. Crick, *The Astonishing Hypothesis: The Scientific Search for the Soul*, Simon & Schuster, New York, NY., 1994.

[6] C. G. Cullen, *Matrices and Linear Transformations*, Dover Publications, Inc., Mineola, NY, 2 ed., 1990.

[7] M. A. Fleischer, *Cybernetic optimization by simulated annealing: Accelerating convergence by parallel processing and probabilistic feedback control*, Journal of Heuristics, 1 (1996).

[8] M. A. Fleischer, *Principles of cybernetics: Potential application to temporally trained neural networks*, Proceedings of the American Society of Engineering Management, (1997).

[9] M. A. Fleischer, *Metaheuristics: Advances and Trends in Local Search Paradigms for Optimization*, Kluwer Academic Publishers, 1999, ch. 28: Generalized Cybernetic Optimization: Solving Continuous Variable Problems, pp. 403–418.

[10] A. Goswami, *The Self-Aware Universe: How Consciousness Creates the Material World*, G. P. Putnam & Sons, New York, NY., 1993.

[11] M. Minsky and S. Papert, *Perceptrons*, MIT Press, Inc., Cambridge, MA., 1969.

[12] A. Sandra and W. J. Coons, *Core Concepts in Embryology*, Lippincott-Raven Publishers, New York, N.Y., 1st ed., January 1998, ch. 6-7, pp. 12–18.

[13] F. Szidarovszky and A. T. Bahill, *Linear Systems Theory*, CRC Press, 1992.

[14] W. A. P. W. S. McCulloch, *A logical calculus of the ideas immanent in neural nets*, Bulletin of Mathematical Biophysics, 5 (1943), p. 115.

[15] N. Wiener, *Cybernetics: Or Control and Communication in the Animal and the Machine*, The M.I.T. Press, Inc., Cambridge, MA., 2 ed., 1961.

[16] S.-S. Yeh and P.-L. Hsu, *Theory and applications of the robust cross-coupled control design*, Journal of Dynamic Systems, Measurement, and Control, 121 (1999), pp. 524–530.