EIGENVALUE ANALYSIS OF DOMINANT FEEDBACK LOOPS

Nathan B. Forrester
Assistant Professor of Management and Economics
College of Business Administration
University of Nebraska - Lincoln
Lincoln, Nebraska 68588

ABSTRACT

Eigenvalue analysis of dominant feedback loops promises to be a powerful new tool for identifying the structural origins of behavior in system dynamics models. Traditional simulation methods for dominant loop analysis are time-consuming and error-prone. A new technique permits calculating the marginal contribution of each feedback loop to each mode of behavior in a model. The technique computes the elasticity of each eigenvalue with respect to the gain of each loop. The elasticities are complex numbers showing the percentage change in natural frequency and damping of each eigenvalue resulting from a one percent change in loop gain. The magnitude of an elasticity measures the overall importance of a loop to a mode of behavior. The magnitudes can be used to rank loops by relative dominance over each mode, or to rank; modes by relative importance to each loop. The technique can be used to analyze both linear and some nonlinear behavior modes.

I. Introduction.

Resolved: The paradigm of system dynamics should be expanded to include eigenvalue analysis of dominant feedback loops.

Identifying the structural origins of behavior in complex system dynamics models is currently a difficult and error-prone process. Identifying dominant feedback loops has traditionally been done by one of two approaches. The first method involves disconnecting unimportant loops and showing that the remaining, isolated loops produce behavior similar to that of the whole model. The second approach involves making small changes in model parameters then observing the changes in model behavior. Loops containing influential parameters are assumed to be dominant.

The advantages and disadvantages of the two traditional approaches are examined in detail below. Both approaches, however, are subject to major shortcomings. Analysis by either method, is subject to error and requires a large number of simulation runs.

A new technique involving eigenvalue analysis of loop dominance overcomes the limitations of both traditional approaches. The eigenvalue method provides exhausive analysis of loop dominance without the need for simulation. The new approach computes the sensitivity of each eigenvalue of a model to a change in the gain of each feedback loop. The analysis can be made completely automatic, eliminating the uncertainties associated with simulation approaches. The amount of computation required is relatively small and complete results can be reported in a two-dimensional table.

Unfortunately the eigenvalue method introduces its own limitations. The new approach can be applied directly only to the analysis of linear behavior modes. However, variants of the

method are discussed that extend its usefulness to at least some of the nonlinear behavior modes commonly encountered in system dynamics models.

II. Two Traditional Approaches to Analyzing Loop Dominance.

The Loop Isolation Approach.

One traditional method for identifying dominant feedback loops involves isolating dominant substructures within a complete model. First, the behavior of the whole model is determined by simulation. Important behavior modes are identified. A hypothesis is then formed about which substructure is responsible for the important modes. The hypoduces the modes of interest, the hypothesis is not refuted. If the isolated substructure does not reproduce the important modes, a new hypothesis is formed and new tests are performed. The end result is a small number of reported tests showing that a subset of all loops can generate behavior modes observed in the complete model.

A substructure can be isolated by blocking the passage of information to or from decoupled parts of the system. Blocking the flow of information is usually accomplished by setting time constants to infinity, setting parameters to zero, flattening table functions, or deactivating switches. If time constants, parameters or tables are not present in the proper information channels, SWITCH functions can be added to permit blockage.

The loop isolation technique has two important advantages:

- The concept is simple and intuitively appealing. If a part of a model reproduces the behavior of the whole, then the part probably contains the feedback loops responsible for the behavior of the complete model.
- The loop isolation approach retains the full nonlinearity of the feedback loops that are isolated. No approximations need be made.

The loop isolation technique is also subject to several limitations:

- A unique time constant, parameter, table, or switch is required in almost every causal link between variables to permit testing of various hypothises about loop dominance. The resulting clutter in model equations is confusing to anyone trying to understand the model structure.
- The number of tests required to examine the behavior of every possible substructure is very large. A separate simulation run is required to test each combination of feedback loops in the model. The maximum number of possible tests is 2^n where n is the number of distinct feedback loops. (The actual number of tests would be somewhat less than 2^n because some of the combinations involve loops that do not connect with each other.) A typical ratio of feedback loops to level variables in a well-structured

system dynamics model may be about three- or five-to-one. A twenty-level model might have some 70 feedback loops. The total number of possible combinations of loops in substructures would be about 1.2×10^{21}. Even after dropping 90% of these combinations because not all the loops interconnect, the remaining number of substructures to be tested is on the order of 10^{20}. Clearly, exhaustive testing by simulation would be prohibitively time-consuming and expensive.
- The loop isolation technique does not identify the relative importance of loops included in the substructure. Some of the included loops may be critical while others might be dropped with minimal consequences. A positive test simply indicates that a sufficient set of loops has been identified, but not that every member of the set is necessary to reproduce important behavior modes of the complete model.
- The loop isolation technique does not indicate the possible importance of loops excluded from the substructure under examination. An excluded loop could have a cancelling effect on an included loop. A behavior mode might be falsely attributed to an included loop while the true cause lay elsewhere in the model. In a similar vein, two excluded two excluded loops might individually have important effects on model behavior but, when taken together, have mutually cancelling effects.

In sum, the loop isolation approach is a simple, intuitively appealing way to find substructures of a model that are capable of creating important behavior modes seen in a complete model. The approach retains all nonlinearities present in the substructure. The method does not guarantee that a particular substructure contains either all the important loops or only the important loops. An exhaustive analysis of loop dominance by the isolation technique is virtually impossible given the enormous number of simulation runs required.

The Parameter Variation Approach.

A second traditional approach to dominant loop analysis involves varying parameters in a model and observing the induced changes in behavior. The first step is to simulate model behavior and identify important modes. The next step is to form a hypothesis about loop dominance, selecting the feedback loop(s) thought to control important modes. The hypothesis is then tested by varying a parameter or table function in the loop(s). The behavior of the model is simulated again with the parameter change. If the important behavior modes are changed significantly, the hypothesis is not refuted. If the behavior of the model is not affected by the parameter change, then the hypothesis must be revised and new tests must be performed. In the final analysis, loops containing parameters with the greatest effects on behavior are reported as dominant.

The parameter variation approach to dominant loop analysis offers several advantages:

- The concept is simple and intuitively appealing. If changing a parameter significantly affects behavior, then the parameter must lie in an important feedback loop.
- The parameter variation approach retains all the nonlinearities present in the model. No approximation or simplification is made to the original nonlinear structure.
- The parameter variation method, unlike the loop isolation approach, does not require deactivating pieces of model structure. The full model is active in every simulation test.
- The parameter variation approach provides a meaure of the relative importance of different loops. Parameters with larger effects on important behavior modes lie in the more important loops.

The parameter variation technique also has major limitations:

- Isolating the effect of a particular loop may not be possible with a single-parameter test. Every parameter in a loop may lie in other loops as well. Without a comprehensive computerized search, it is difficult to ascertain how many feedback loops involve a particular parameter or table. A proper test might be constructed by varying more than one parameter at a time. Offsetting effects from multiple parameter changes would leave only one loop altered. Such tests, however, are very difficult to construct.

- To insure that no important loop is overlooked, a separate simulation test must be performed for every loop in the model. A twenty-level model might require 50 to 100 simulation runs for complete testing.
- The effect of a parameter change on a particular behavior mode is often difficult to determine from simulation output. A simulation run shows the effect of a particular change on all modes simultaneously. Multiple modes are difficult to separate visually in plotted output. The need for visual interpretation invites error.

In sum, the parameter variation approach to dominant loop analysis is a simple and intuitively appealing way to assess the marginal contribution of feedback loops to the various behavior modes of a model. The approach retains the the full model structure including all nonlinearities at all stages of analysis. The technique does not, however, guarantee finding all important loops unless every loop is identified and tested separately. Tests are difficult to construct when parameters and tables lie in more than one loop, as they do in most system dynamics models. Furthermore, the effects of a loop on different behavior modes are difficult to untangle.

III. Dominant Loop Identification Via Eigenvalue Analysis.

A new approach to identifying dominant loops is proposed here. The new method involves eigenvalue analysis of model behavior. Eigenvalues correspond to the linear behavior modes present in a model. Eigenvalues can be computed directly from knowledge of system structure and parameters without need for simulation. The technique computes the sensitivity of the eigenvalues (behavior modes) to changes in the strength of each feedback loop. An intermediate result is an array of complex numbers measuring the effect of a change in the gain of each loop on the freqcannot currently be applied directly to nonlinear systems. Linearization, however, may not be as severe a limitation as first imagined. The next section explores how the eigenvalue approach can be used to understand the origins of some nonlinear modes encountered in system dynamics models. Linearization requires that each variable be expressed as a linear combination of other variables in the model:

$$V_i = \sum^{j} a_{ij} \cdot V_j \qquad (1)$$

where each constant parameter a_{ij} is the partial derivative of variable V_i with respect to variable V_j:

$$a_{ij} = \frac{\partial V_i}{\partial V_j} \qquad (2)$$

The derivatives can be calculated either analytically or experimentally by vaying the value of V_j and observing the change in V_i. Each coefficient a_{ij} is referred to as the gain of the causal link between variable V_j and variable V_i.

The individual equations for each variable are combined into a system of equations describing the entire model:

$$\begin{bmatrix} \dot{\underline{X}} \\ \underline{Y} \end{bmatrix} = \begin{bmatrix} \underline{A}^*_{11} & \underline{A}^*_{12} \\ \underline{A}^*_{21} & \underline{A}^*_{22} \end{bmatrix} \begin{bmatrix} \underline{X} \\ \underline{Y} \end{bmatrix} \qquad (3)$$

where:

$\dot{\underline{X}}$ - vector of level variable derivatives (net rates)

$\underline{X}$ - vector of level variables

$\underline{Y}$ - vector of non-level variables (rates and auxiliaries)

$\underline{A}^*$ - matrix of partial derivatives a_{ij} from (1).

For eigenvalue analysis, equation (3) must be collapsed to form:

$$\dot{\underline{X}} = \underline{A} \quad \underline{X} \qquad (4)$$

The matrix A in equation (4) is computed from A^* in equation (3) by the formula:

$$\underline{A} = \underline{A}^*_{11} + \underline{A}^*_{12} \left[\underline{I} - \underline{A}^*_{22}\right]^{-1} \underline{A}^*_{21} \qquad (5)$$

The eigenvalues of matrix A are then computed. The eigenvalues describe all the linear behavior modes preseent in the model. Eigenvalues are, in general, complex numbers that can be plotted in the complex plane. Figure 1 shows several eigenvalues plotted in the complex plane and the corresponding dynamic behavior patterns they represent. An eigenvalue with no imaginary part corresponds to a mode of exponential growth or decay. Eigenvalues with non-zero imaginary parts come in complex conjugate pairs and correspond to oscillatory modes of behavior. [1]

The next step is to compute the sensitivity of each eigenvalue to a change in the strength of each feedback loop. The sensitivity measure is based on the partial derivative of the eigenvalue with respect to the gain of the loop. Figure 2 shows the partial derivative of an eigenvalue with respect to the gain of a loop plotted as an extension of the eigenvalue vector. The partial derivative shows the absolute change in the eigenvalue for a change in loop gain.

The raw partial derivative is then multiplied by the loop gain and divided by the eigenvalue to form a dimensionless index of sensitivity. The index is called the eigenvalue elasticity with respect to loop gain, or loop gain elasticity for short. The elasticity is the percentage change in the eigenvalue resulting from a one percent change in the gain of the loop. The elasticity is computed as:

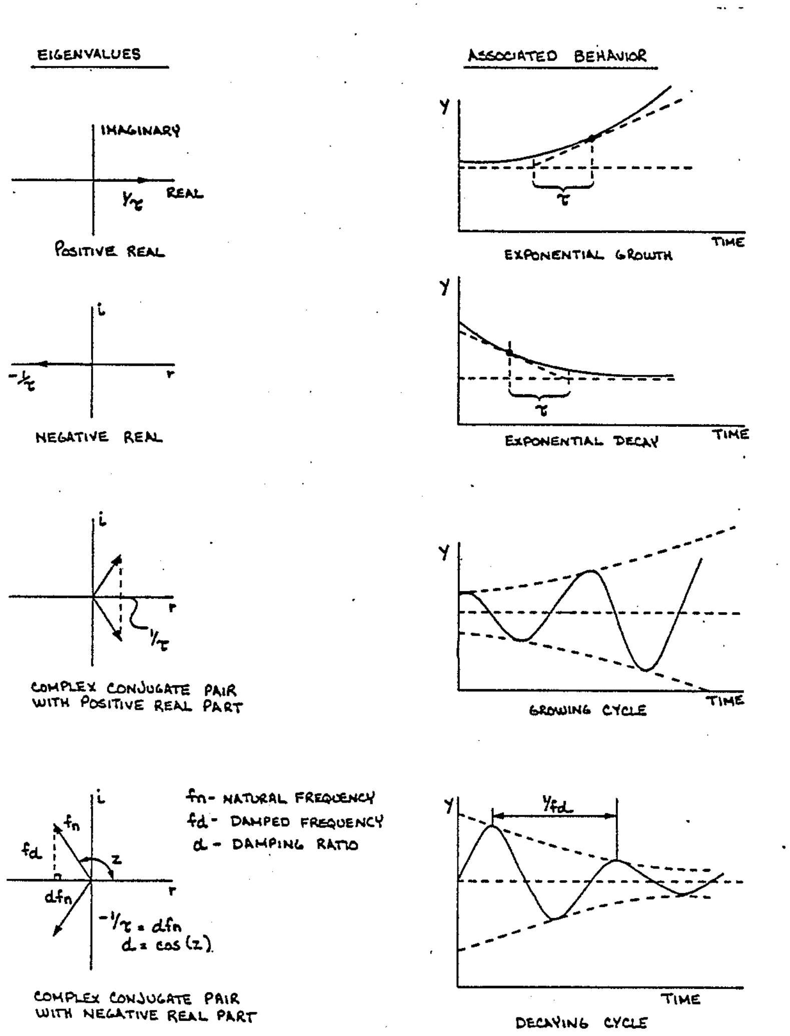

Figure 1. Typical Eigenvalues and Associated Behavior Modes

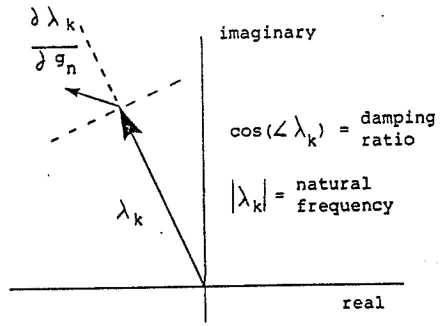

Figure 2. Eigenvalue Sensitivity

$$E_{kn} = \frac{\partial \lambda_k}{\partial g_n} \cdot \frac{g_n}{\lambda_k} \qquad (6)$$

where:

E_{kn} - elasticity of the kth eigenvalue with respect to the gain of the nth loop

λ_k - kth eigenvalue

g_n - gain of the nth loop.

Figure 3 shows the eigenvalue elasticity corresponding to the partial derivative shown in Figure 2. Multiplication by g_n/λ_k rotates and scales the dotted axes in Figure 2 so that the partial detrivative refers to a standardized eigenvalue of length one and angle zero. The real component of the elasticity then shows the effect of loop gain on the natural frequency of the eigenvalue. The imaginary component of the elasticity shows the effect of loop gain on the damping ratio of the eigenvalue. [2]

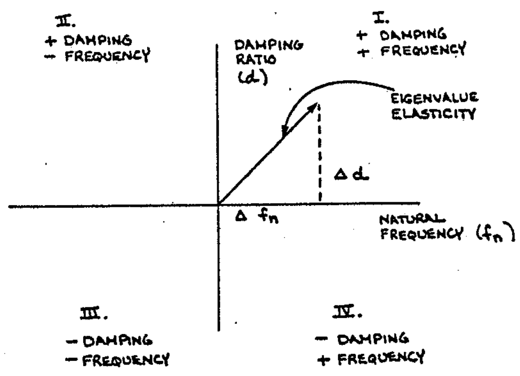

Figure 3. Loop Gain Elasticity

The eigenvalue elasticities with respect to loop gains are not computed directly. Instead, the loop gain elasticities are computed indirectly from the elasticiites of the eigenvalues with respect to the gains of the individual causal links between variables. The elasticity of the kth eigenvalue with respect to the gain of the causal link between variable V_j and variable V_i is defined as:

$$E_{kij} = \frac{\partial \lambda_k}{\partial a_{ij}} \cdot \frac{a_{ij}}{\lambda_k} \tag{7}$$

and computed according to the formula: [3]

$$E_{kij} = \underline{L}_k' \frac{\partial \underline{A}}{\partial a_{ij}} \underline{R}_k \cdot \frac{a_{ij}}{\lambda_k} \tag{8}$$

where $\underline{L}_k$ and $\underline{R}_k$ are the left and right eigenvectors associated with the kth eigenvalue. The right eigenvectors are ususally

computed by packaged programs that compute eigenvalues. The matrix of left eigenvectors can be computed as the inverse of the martrix of right eigenvectors. The partial derivative of the matrix A with respect to the gain of the link between V_j and V_i can be computed by reevaluating equation (5) before and after a small change in the value of the gain a_{ij}. The change in matrix A is divided by the change in a_{ij} to obtain the partial derivative:

$$\frac{\partial A}{\partial a_{ij}} = \frac{\Delta A}{\Delta a_{ij}} = \frac{\Delta\left[\underline{A}_{11}^* + \underline{A}_{12}^* (\underline{I} - \underline{A}_{22}^*)^{-1} \underline{A}_{21}^*\right]}{\Delta a_{ij}} \tag{9}$$

The eigenvalue elasticities with respect to loop gains can now be determined from the elasticities with respect to the link gains. By Mason's Rule the coefficients of the characteristic equation of a system can be expressed as functions of loop gains. Because the eigenvalues of a system are determined by the coefficients of the characteristic equation, the eigenvalues are also functions of the loop gains. Therefore, the link gain elasticity formula in (7) can be rewritten as:

$$E_{kij} = \frac{\partial \lambda k}{\partial a_{ij}} \cdot \frac{a_{ij}}{\lambda_k} = \frac{\partial \lambda k}{\partial g_n} \cdot \frac{\partial g_n}{\partial a_{ij}} \cdot \frac{a_{ij}}{\lambda_k} \tag{10}$$

where g_n is the gain of loop n. Because the gain of a loop is equal to the product of gains of the causal links in the loop,

$$\frac{\partial g_n}{\partial a_{ij}} = 0 \tag{11}$$

when the link from variable V_j to V_i does not lie in loop n; and

$$\frac{\partial g_n}{\partial a_{ij}} = \frac{g_n}{a_{ij}} \qquad (12)$$

when the link from V_j to V_i does lie in the loop. In light of (11) and (12) above, equation (7) can be rewritten as:

$$E_{kij} = \sum^{*} \frac{\partial \lambda_k}{\partial g_n} \cdot \frac{g_n}{\lambda_k} \qquad (13)$$

where * indicates summation over all loops containing the link from V_j to V_i. In other words, the elasticity of an eigenvalue with respect to the gain of a link is equal to the sum of the elasticities of the eigenvalue with respect to the gains of all the loops that pass through the link.[1]

Result (13) permits setting up a system of simultaneous linear equations for each eigenvalue. The equations state that each link gain elasticity equals the sum of the loop gain elasticities of every loop passing through the link. Setting up the equations requires that all feedback loops in the model be identified. An automatic loop-search algorythm can be used. A feedback loop is defined here in the conventional manner as a circular chain of causal links that does not contain the same variable more than once.

1. Use of Mason's Rule to demonstrate result (13) was suggested by Stanley M. Liberty at the University of Nebraska - Lincoln.

After finding all loops and the links contained in each loop, the systems of simultaneous equations can be written. Having already evaluated the link gain elasticities, the equations can be solved for the loop gain elasticities. In general, the equations will over-determine the loop gain elasticities (there will be more links than loops in a typical model) but the equations will not be inconsistent. After solving the system of equations for each eigenvalue, the results can be compiled into a two-dimensional array relating eigenvalues to loops in the model. Each entry in the array is the elasticity of an eigenvalue with respect to the gain of a loop. The entries are complex numbers showing the percentage change in both frequency and damping of the eigenvalue for a one percent change in loop gain. The magnitude of each complex number gives an overall index of the importance of a feedback loop to an eigenvalue. The magnitudes can be used to generate a list ranking the feedback loops in order of relative importance to a particular eigenvalue. The magnitudes can also be used to create a list of eigenvalues in order of relative importance for a particular loop. The raw elasticities can be used to determine how an "important" loop contributes to the natural frequency and damping ratio of an eigenvalue.

Applications of the Eigenvalue Approach.

As seen in the preceding section, the eigenvalue approach to dominant loop analysis measures directly the marginal contribution of every feedback loop to every behavior mode in the linear approximation to an underlying nonlinear model. The method permits ranking of the feedback loops by importance to a given eigenvalue and/or ranking of the eigenvalues by importance for a given loop. The method can be used directly to analyze the behavior of linear models or quasi-linear models where the role of nonlinearities is relatively minor. The approach can also be used directly to analyze linear behavior modes within strongly nonlinear models. The method can be used indirectly to analyze the structural origin of certain nonlinear behavior modes as well.

The linear behavior modes amenable to direct analysis by the eigenvalue technique are:

-Exponential growth,

-Exponential decay,

-Damped oscillation,

-Exploding oscillation, and

-Combinations of the above.

The nonlinear behavior modes amenable to indirect analysis by the eigenvalue approach include:

-"S"-shaped growth,

-Overshoot-and-collapse, and

-Limit cycles.

To analyze nonlinear modes of S-shaped growth and overshoot-and-collapse, the eigenvalue method can be applied repeatedly at different operating points during the transition phases from growth-to-stagnation, stagnation-to-decline, and decline-to equilibrium. The results must be interpreted carefully. Both the eigenvalues and the loops that dominate them will change during the transition periods. By carefully tracking the chnages, the shifting loop dominance that creates the nonlinear behavior modes can be ascertained.

To analyze the dominant loops in a limit cycle, the eigenvalue approach can be applied directly at the underlying unstable equilibrium point to determine the causes of the explosive cycle. The analysis can be used again at the extreme swings of the cycle to determine which loops keep the amplitude of the cycle from growing indefinitely.

The practicality of using eigenvalue analysis of loop dominance in nonlinear models has not yet been determined. While the approach appears to show promise, a software package has not yet been developed to permit experimentation. The approach may prove to be too costly to apply repeatedly. The interpretation of shifting eigenvalues and loop dominance may prove too difficult.

Advantages of the Eigenvalue Approach.

The eigenvalue approach to analyzing loop dominance has several advantages over the two traditional approaches examined at the beginning of this paper.

- The eigenvalue technique directly measures the impact of loops on behavior rather than the impact of parameters on behavior.
- The technique can provide a one-dimensional measure of dominance for ranking loops by their importance to a mode.
- The eigenvalue approach automatically provides exhaustive analysis of the impact of every loop on every mode.
- The technique offers great saving both in the amount of computation required to obtain results and in the amount of time required to interpret those results.
- The technique offers a standardized, reproducible procedure for analyzing loop dominance.
- The complete results of eigenvalue analysis can be presented in a very compact form.

Limitations of the Eigenvalue Approach.

The eigenvalue technique proposed here is certainly not without limitations of its own. Some of those limitations are inherent in the theory and are examined below. Other limitations may become apparent when the technique is actually applied. The

limitations currently known are:

- The technique measures only the change in eigenvalues induced by small incremental changes in the gain of each feedback loop. The loops judged to be dominant by this criterion may not be capable of reproducing the mode in question when isolated from the rest of the model structure. The loop dominance analysis may not be a good guide to model simplification.
- Results of the eigenvalue analysis are only strictly applicable to the linear modes around a particular operating point in nonlinear systems. The analysis must be applied repeatedly at different operating points to analyze the shifting loop dominance that accounts for nonlinear behavior modes. The necessary interpretation of results at different operating points opens the door to errors in analysis. Only experience will tell whether the technique is useful for analyzing the nonlinear modes commonly encountered in system dynamics models.
- No software is currently available to permit application of the eigenvalue technique to system dynamics models. At least two efforts are in progress but no dates have yet been set for shipment of a user-friendly software packages.

Extensions of the Eigenvalue Approach.

Several possible extensions of the current work should be noted:

- The approach can be extended to include techniques for analyzing nonlinear behavior modes. In its simplest form such an extension might be a set of rules-of-thumb for interpretation of results from reapplication of the methods outlined here at different operating points. A more sophisticated extension might use similar techniques to examine the behavior of higher-order approximations to a full nonlinear model. The linear analysis examines only the properties of the first term in a Taylor series expansion of a nonlinear system. The properties of the higher-order terms may also yield to a qualitatively similar form of analysis.
- The eigenvalue approach could be modified to examine the dominance of feedback loops in creating the frequency response characteristics of a system. Frequency repsonse may be a better way to measure behavior in oscillatory systems than eigenvalues.
- The eigenvalue analysis of dominant loops, while designed for another purpose, might be a useful guide to reducing model complexity. Loops with a small marginal contribution to a given eigenvalue might be eliminated without significantly altering the behavior mode. Neither the theoretical nor empirical implications of such use have been examined.

IV. Conclusions.

Two techniques are traditionally used to determine which feedback loops are responsible for behavior in a system dynamics model. Both techniques have major drawbacks. The first technique involves isolating a set of loops from the rest of the model to show that they can produce the behavior mode in question. The loop isolation approach does not indicate whether or not all loops in the set are important to the mode, nor does it indicate the existence of important loops not in the selected set. The second technique involves changing the parameters and tables of a model and observing the effects on behavior. The parameter variation approach is cumbersome because a single parameter change may alter several loops simultaneously; the effects of an individual loop are difficult to measure. The parameter variation approach is further complicated by the fact that a parameter change may affect several different behavior modes; the effects on different modes are difficult to untangle from simulation output. Both traditional approaches require an extraordinarily large number of simulation runs for complete analysis of loop dominance.

A new approach to analyzing loop dominance is suggested here. The new technique involves linearizing the underlying nonlinear simulation model. Then the eigenvalues of the linear approximation are computed. Each eigenvalue corresponds to a mode of behavior in the model. The sensitivity of each eigenvalue to a change in the gain of each feedback loop is then calculated. The sensitivity is expressed as an elasticity, a dimensionless ratio, showing the percentage change in the eigenvalue for a one percent change in the gain of the loop. The elasticity is a complex number whose real and imaginary parts show the effects of a change in loop gain on the natural frequency and damping ratio of the eigenvalue respectively. Results of the analysis can be reported as a two-dimensional array of elasticities relating each eigenvalue to each loop. The array can be used to generate a ranking of feedback loops by importance to each eigenvalue or a ranking of eigenvalues by importance for each loop. The magnitude of the elasticities can be used as a one-dimensional measure of "importance" for the purpose of ranking.

The new technique prodives a compact, automatic, reproducible method for finding dominant loops. The computational burden is small relative to the traditional approaches. The analysis is also exhaustive; the effect of every loop on every mode is calculated.

The new technique is directly applicable linear and nearly-linear systems and can also be used to analyze linear modes within nonlinear systems. The technique may be applied indirectly to the analysis of nonlinear behavior modes. The

practical utility of the technique for analyzing the nonlinear modes commonly encounterd in system dynamics models has not yet been determined.

Unfortunately, no user-friendly software package is currently available that permits eigenvalue analysis of loop dominance. Prototype packages should be available within a year. When the necessary software becomes avalable, eigenvalue analysis of loop dominance should become an important new tool for practitioners of system dynamics.

Referrences:

1 . Ogata, Katsuhiko. Modern Control Engineering, Englewood Cliffs, N.J.: Prentice Hall, Inc., 1970.

2 . Forrester, Nathan B. A Dynamic Synthesis of Basic Macroeconomic Theory: Implications for Stabilization Policy Analysis. PhD. Dissertation in Management, M.I.T., 1982.

3 . Porter, B. and R. Crossley. Modal Control Theory and Applications. Taylor and Francis, 1972.