

THESIS

22403229



This is to certify that the

dissertation entitled

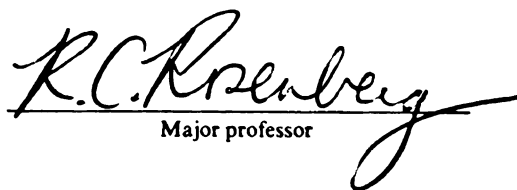
Improving Design Assessment and Simulation
of Large-Scale Dynamic Systems

presented by

Tong Zhou

has been accepted towards fulfillment
of the requirements for

Ph.D degree in Engineering


Major professor

Date 8/12/88

LIBRARY
Michigan State
University



RETURNING MATERIALS:
Place in book drop to
remove this checkout from
your record. FINES will
be charged if book is
returned after the date
stamped below.

FEB 9 / 1984
229

IMPROVING DESIGN ASSESSMENT AND SIMULATION
OF LARGE-SCALE DYNAMIC SYSTEMS

By

Tong Zhou

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Mechanical Engineering

1988

ABSTRACT

IMPROVING DESIGN ASSESSMENT AND SIMULATION
OF LARGE-SCALE DYNAMIC SYSTEMS

By

Tong Zhou

When accurate models of complex large-scale engineering systems are made, they may involve large numbers of coupled nonlinear differential and algebraic equations. It can be both time-consuming and difficult for the design engineers to get a good assessment of the performance of the design, especially in a way leading to design improvements. The objective of this research is to find methods that improve the design assessment and simulation of complex, large-scale dynamic models, thereby reducing time spent to gain increasing insight by the engineers.

With respect to simulation efficiency, the problem of implicit R fields is very important to the bond graph modeling of engineering systems. Simulation efficiency in such problems may be increased by reducing the number of iteration variables. A new algorithm for determining the minimum number of iteration variables required in a model with implicit R fields is presented. The algorithm generates exact results for implicit R fields containing 1-port R nodes and weighted junction structures (1, 0, TF). An extension is made for causal IRFs containing multiport R nodes to get the minimum number of iteration

variables. As a further extension the basis order properties of general junction structures (1, 0, TF, GY) were derived by using gyrographs and a maximum matching algorithm.

To increase insight a new approach to model order reduction and simplification has been developed in the bond graph framework. A pilot version has been implemented in software. Power responses measured in various ways are displayed directly on the bond graph by color coding. Such a display gives engineers an easy way to see the power distributions in a large-scale dynamic system under various sets of operating conditions. By interpreting this power-based data suitably, possibilities for model order reduction and simplification can be identified without detailed analysis or equation manipulation. The power response data is available for nonlinear models as well as for linear ones, so the power-based reduction approach may be used for nonlinear time-varying systems.

The results obtained contribute to the design assessment and simulation problem by increasing simulation efficiency for certain classes of models and by increasing the insight about system behavior available to the design engineers. Further directions of work are discussed based on the results presented.

Dedicated to

My parents, Wenjin Zhou, MD., Hangxiang Lee, MD..

My wife Shuling Jiang. and

My son Fan

ACKNOWLEDGMENTS

The author is very grateful to Dr. Ronald C. Rosenberg for his guidance, encouragement, persistence, and friendship in his role as major advisor, critic, taskmaster, and friend. The author also would like to thank Dr. Hasan Khalil, Dr. Clark Radcliffe, and Dr. Steve Dragosh for their contribution throughout his graduate study and research in both class teaching and serving on his doctoral committee. Many thanks go to Mr. Dave Withers of the IBM Information Systems for his interest and support in the research and for serving on the author's guidance committee.

The support of the IBM Thomas J. Watson Research Center is greatly appreciated. The research was also supported by the use of the excellent computing facilities and the aid of the cooperative personnel at the A. H. Case Center for Computer-aided Engineering in the College of Engineering at Michigan State University.

TABLE OF CONTENTS

LIST OF TABLES	ix
LIST OF FIGURES	x
Chapter 1 INTRODUCTION	1
1.1 Problem Statement	1
1.2 Research Objectives	2
1.3 Dissertation Organization	4
Chapter 2 EFFICIENT COMPUTATION OF IMPLICIT DISSIPATION FIELDS ...	6
2.1 Problem Definition	6
2.2 General Solution by Iterative Methods	12
2.3 IRFs with One-Port R Elements.....	19
2.4 IRFs with Multiport R Elements.....	25
2.4.1 Multiport resistances in physical systems	25
2.4.2 Extension of the Algorithm to IRFs with multiport R nodes	28
2.4.3 Assigning causality to an IRF Containing One Multiport R Node	42
2.5 IRFs with General Junction Structure	58
2.5.1 General Junction Structure	58
2.5.2 Gyrographs and Maximum Matching	59
2.5.3 Basis Order Algorithm for General Junction Structures	66
2.5.4 Examples	74
2.5.5 Remarks on the Algorithm	77
Chapter 3 MODEL ORDER REDUCTION METHODS	79
3.1 Aggregation Methods	80
3.2 Modal Analysis	82
3.3 Lyapunov Function Method	85

3.4 Perturbation Methods	86
3.4.1 Regular perturbation method	86
3.4.2 Singular perturbation method	87
3.5 Component Connection Methods	89
3.5.1 Component connection method	89
3.5.2 Diakoptics	91
3.6 Component Cost Analysis of Large Scale Systems	92
3.7 Model Order Reduction in Bond Graph Models	93
3.7.1 Sequential causality assignment method	94
3.7.2 Reciprocal system method	95
3.7.3 Finite mode distributed system models	96
Chapter 4 POWER PORTRAITS OF DYNAMIC SYSTEM MODELS	98
4.1 Graphical Representations of Dynamic System Models.....	98
4.2 Uniform Performance Measure --- Power	101
4.3 Power/energy Visualization	105
4.3.1 Computation of power variables	105
4.3.2 Power measures	110
4.3.3 Scaling and color-coding	113
4.4 Power Method for Model Order Reduction and Simplification	113
4.4.1 Simplification of physical effects	114
4.4.2 Decomposition and simplification of subsystems	116
4.4.3 Procedure for model simplification	117
4.5 Examples	118
4.5.1 A radar pedestal unit	118
4.5.2 An Euler-Bernoulli beam with a vibratory load	127
Chapter 5 SUMMARY AND CONCLUSIONS	137
5.1 Implicit R-field Simulation	137
5.2 Power Distribution and Display	139

APPENDIX

Appendix	Standard Bond Graphs and Gyrobondgraphs	142
	1. Standard bond graphs	
	2. Gyrobondgraphs	
	3. Gyrographs	

BIBLIOGRAPHY		145
--------------	--	-----

LIST OF TABLES

Table	page
2-1 Possible variable selection	38
4-1 Powers in various energy domains	104
4-2 Parameters of pedestal model	121
4-3 Effect indices	122
4-4 Eigenvalues of the original and reduced models	125
4-5 The power listing of five-mode model	130
4-6 The powers on cut bonds	132
4-7 The power listing of two-mode model	134
4-8 Eigenvalues of the original and reduced models	134
4-9 Computer times	134
A-1 Standard node sets	144

LIST OF FIGURES

Figure	page
2-1 A mechanical system with dissipative coupling	7
(a) Physical system	
(b) Bond graph model	
(c) Causality assignment	
2-2 Multiport Field Structure	13
2-3 Key vectors	13
2-4 A system containing two IRFs	18
(a) The identified IRFs	
(b) Causality of the IRFs	
2-5 An example of an IRF	24
(a) Bond graph of the R-field	
(b) Causality assignment	
2-6 Four-way-valve amplifier	26
2-7 Reduced form of hydraulic amplifier	26
2-8 Power transistor and its bond graph model	27
2-9 Key vectors in an IRF	28
2-10 Functional diagram of an IRF	32
2-11 Variable flow diagram	33
2-12 Loop formed by an improper vortex exchange	36
2-13 Inclusion of 1-port R nodes	37
2-14 An electronic circuit and its bond graph model	39
2-15 Acausal IRF with multiport R node	42
2-16 An example of self-loop and R-block	43
2-17 An R-block of type 1	45
2-18 An R-block of type 2	46
2-19 An R-block of type 3	46
2-20 An R-block of type 4	47
2-21 A multiport R node without structural coupling	51

2-22	Example 1	54
2-23	Example 2	56
2-24	Graphs of an R-field	62
	(a) The standard bond graph	
	(b) The gyrobondgraph	
	(c) The gyrograph	
2-25	Causality in different graphs	64
	(a) In a standard bond graph	
	(b) In a gyrobondgraph	
	(c) In a gyrograph	
2-26	A graph and one of its matchings	66
2-27	An alternating path	69
	(a) Before reversing markings	
	(b) After reversing markings	
2-28	An alternating cycle	69
	(a) Before reversing markings	
	(b) After reversing markings	
2-29	Example 1	75
	(a) Bond graph and gyrobondgraph	
	(b) Gyrograph	
	(c) Maximum matching in G_J	
	(d) Maximum matching in GG	
	(e) Causality corresponding to the F-minimum basis	
	(f) Causality corresponding to the F-maximum basis	
2-30	Example 2	76
	(a) Gyrograph	
	(b) Maximum matching of G_J	
	(c) Maximum matching of GG	
3-1	Reciprocal system method	95
3-2	Finite modal bond graph model	97
4-1	Graphical representations of mathematical modeling	99
4-2	Powers in system models	103
	(a) schematic diagram	
	(b) block diagram	
	(c) bond graph	
4-3	Power responses of the system in Figure 4-2	111
4-4	Powers on external bonds	115
4-5	Powers on cut bonds	117

4-6	A radar pedestal unit	120
4-7	Bond graph for the radar pedestal unit	120
4-8	Power display on the bond graph of the pedestal	124
4-9	Reduced-order model of the pedestal	125
4-10	Responses of the original and reduced models	126
4-11	A beam-coupled system	128
4-12	Modal bond graph model for the beam coupled system	128
4-13	Power distribution in the modal bond graph model	131
4-14	Reduced model with two modes	133
4-15	Comparison of input and load velocities	135

Chapter 1

INTRODUCTION

1.1. Problem Statement

Suppose we have an energy-based structural model (i.e., a bond graph model) of a large-scale, complex, nonlinear electro-mechanical system, including its controls. In addition we have a set of nominal constitutive equations for the energy/power elements and a set of input functions for the source elements. A simulation can be run to establish the nominal system response under the conditions noted. We assume that the response so obtained is satisfactory in terms of required system performance. This situation represents the nominal design (Rosenberg, 1985).

Now we ask the following question: How much change can be tolerated in a set of parameters associated with the constitutive equations and still have the system response meet the performance criterion acceptably? It is necessary to define an acceptable performance criterion C quantitatively to attack the question meaningfully. We assume this has been done in terms of a set of system output variables $y(t)$.

The most obvious method of attack is to choose a set of values for a parameter vector P , run a simulation, and assess the resulting $y(t)$ with respect to the performance criterion C . There is a statistical version of the problem in which statistics are assigned to the P vector, and the statistics of $y(t)$ with respect to C are generated. The

statistical approach is of great value in making cost/benefit analyses associated with manufacturing and assembly decisions (Prakash, et. al, 1985).

One deterministic version of the problem is based on worst-case analysis. We try to find the limits on the values for P that push the response $y(t)$ to the acceptable limits of C . In another version we try to guarantee that choosing P values with certain bounds will meet C acceptably. The first version is the best we can hope to do with a deterministic approach, given C in an acceptable/unacceptable sense (as contrasted with a cost function sense). Since setting error tolerance on performance of many classes of electro-mechanical equipment is a common practice, we will assume that C is stated in an acceptable/unacceptable sense.

Progress in treating the problem described above will benefit engineers in discovering and eliminating failure modes characterized by exceeding dynamic tolerances within the design cycle. Therefore we can expect to achieve greater reliability and manufacturability of new products.

1.2 Research Objectives

The particular research objectives are derived from the problem described above. Given a fairly general class of nonlinear dynamic systems that are physically (i.e., energetically) based, we expect to use a numerically-implemented approach, rather than functional analysis, as the principal tool in improving practical assessment techniques. Included in numerical computation are tools for studying the effects of

variations in model topology as well as parameter set values on the response $y(t)$ relative to C . The principal goal is to find ways both to reduce computational effort relative to deterministic modeling and to guide the search for model alternatives automatically.

The main problem of interest, which is related to the efficiency of simulation, but more concerned with the solution of the system equations, is that of coupled nonlinear algebraic equations that arise when coupled bond graph R elements exist. The implicit (coupled) nonlinear algebraic equations are often difficult to solve, and they typically must be solved iteratively at each integration step. The primary objectives in this sub-problem are:

1. to develop an automatable method to identify and solve the coupled nonlinear algebraic equations correctly; and
2. to study possible improvements in the existing solution process such that the computation cost can be reduced.

In order to make the design process more efficient and computationally economical, one avenue of approach is to reduce the order of the system model, since the computation effort typically increases much faster than the system dimension does, and high-dimension systems lead to more complicated control design. Since no general theoretical method is available for assessing the response of nonlinear, large-scale complex engineering systems, a need arose to find a new method to consider the possibility of reducing the order of the dynamic system model. One of the research objectives is to provide a general purpose numerical approach to model order reduction and simplification. The application of this approach should not be restricted by the

linearity of the dynamic systems. Also, it should be easy to use and should provide engineers additional information about the system under study.

1.3 Dissertation Organization

In pursuit of these objectives the following subjects will be discussed in the thesis. The dissertation consists of two parts: improved solution methods for implicit dissipation fields and power portraits of dynamic system models. Although both problems are related to the improvement of simulation and design efficiency, they are rather different in the nature. Therefore, we will discuss them separately in the following chapters.

Chapter 2 deals with the implicit dissipation field problem. It presents results concerning the efficient computation of the coupled algebraic equations by an iteration process. A procedure for determining the minimum number of (independent) iteration variables for implicit R fields with simple and weighted junction structures is given. For implicit R fields with general junction structures an algorithm for finding the minimum and maximum flow input variables is described.

In Chapter 3 a brief review of some recent work in the area of model order reduction is given. This chapter provides a background relative to the general model order reduction problem and analytically-oriented methods.

Chapter 4 introduces the power concept and its application in bond graph models. A new tool for exploiting the power attributes of

a dynamic system is presented. Some examples are given to illustrate the potential scope and utility of this method.

In Chapter 5 the results for improving the design assessment tools are summarized and some future research directions are identified.

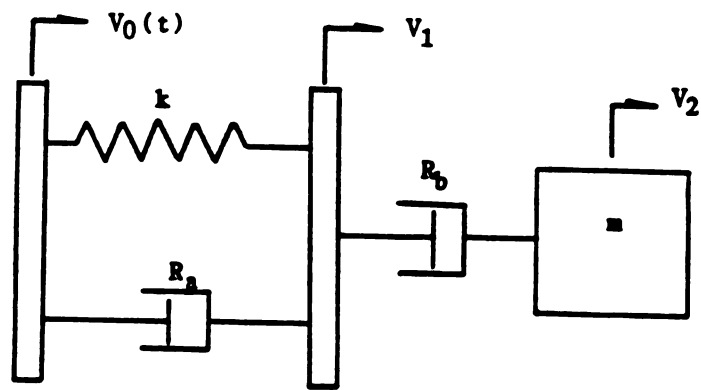
The computer simulations used in this research were run using the ENPORT-7 bond graph/block diagram software. The program was version 7.1 run on a PRIME 750 under PRIMOS. In several cases noted in the text the algorithms developed in this research were implemented within the ENPORT framework on a pilot basis.

EFFICIENT COMPUTATION OF IMPLICIT DISSIPATION FIELDS

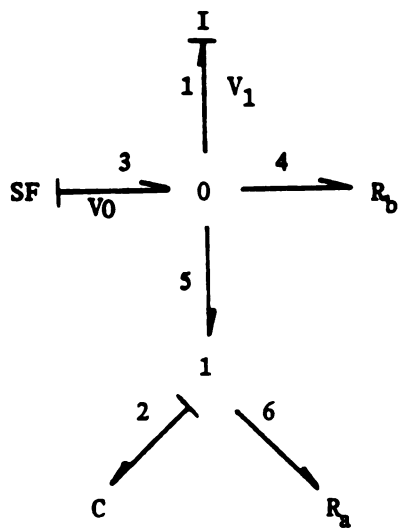
2.1 Problem Definition

One of the major objectives of improving design assessment is to automate the process of formulating and solving the state equations associated with bond graph models of engineering systems (See Appendix A). A second objective is to provide timely and insightful feedback to the designer. There are several sources of difficulty in accomplishing the major objective. The one we wish to focus on here is that of solving coupled nonlinear algebraic equations that arise when the R elements are connected in particular ways in the model. The implicit equations are often difficult to solve, and they typically must be solved several times in each integration step. Consequently, it is helpful to be able to inform the modeler in detail of the existence of such coupling. Furthermore, increasing the efficiency with which such solutions are obtained can dramatically decrease the overall solution time.

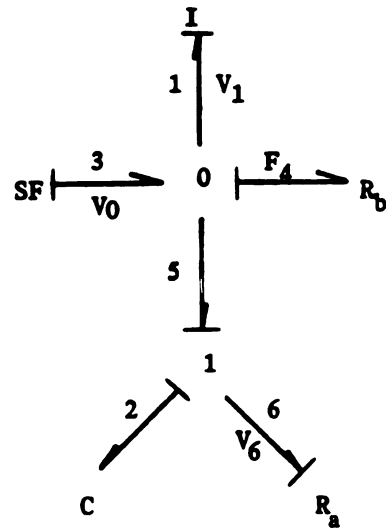
The existence of algebraic loops in the equations of a physical system may not be detected until the sorting or reducing process starts in most traditional simulation approaches. But their existence can be verified even before equation formulation when the bond graph approach is employed. To illustrate the case let us first consider a physical device shown in Figure 2-1a. In Figure 2-1b, the corresponding bond graph model has been built. The I element represents the inertial effect and the compliance effect is indicated by a C element in the mechanical system. The R elements represent energy dissipative effects. The SF



(a)



(b)



(c)

Figure 2-1 A mechanical system with dissipative coupling

(a) Physical system

(b) Bond graph model

(c) Causality assignment

element indicates an imposed velocity on the left plate (assumed massless) as an input. It is desired that the state-space equations be obtained in an explicit form as follows

$$\dot{p}_1 = g_1(p_1, q_2, v_0) \quad (2.1)$$

$$\dot{q}_2 = g_2(p_1, q_2, v_0) \quad (2.2)$$

where p_1 is the momentum of mass I and q_2 is the deflection of spring C.

Causality can be assigned to the bond graph of Figure 2-1b according to the general rules (Rosenberg and Karnopp, 1983). After finishing the first step (assigning required causality to the source SF) and the second step (assigning the integral causality to the storage elements C and I), we find that the causality does not fully extend through the graph. Some acausal bonds (bonds 4, 5, and 6) will be left. At this stage, we realize that an R-field exists in this system. This implies that there will be an algebraic loop in the system equations.

Suppose we continue the causality assignment by imposing an arbitrary causal orientation on one of the two R elements, say, R_4 . Then we extend the causal implication through the graph using the constraint elements (0 and 1). Now the causality assignment has been completed (Figure 2-1c). The state vector X and input vector U are identified as follows

$$X = \begin{bmatrix} p_1 \\ q_2 \end{bmatrix} \quad U = [v_0]$$

If we define F_4 and V_6 as auxiliary variables, then the auxiliary equations are readily obtained as

$$\dot{p}_1 = F_4 \quad (2.3a)$$

$$\dot{q}_2 = V_6 \quad (2.3b)$$

and the constitutive equations are

$$F_4 = g_4(V_4) = g_4(V_0 - p_1/m_1 - V_6) \quad (2.4a)$$

$$V_6 = g_6(F_6) = g_6(F_4 - kq_2) \quad (2.4b)$$

Assume that both R_4 and R_6 are linear, that is,

$$F_4 = R_4 V_4 \quad (2.5a)$$

$$V_6 = R^{-1} F_6 \quad (2.5b)$$

After some algebraic manipulations to eliminate the auxiliary variables F_4 and V_6 , an explicit state-space equation set can be developed; namely,

$$\dot{p}_1 = -[R_4 R_6 / (R_4 + R_6) m] p_1 + [R_4 k / (R_4 + R_6)] q_2 + [R_4 R_6 / (R_4 + R_6)] V_0 \quad (2.6a)$$

$$\dot{q}_2 = -[R_4 / (R_4 + R_6) m] p_1 - [k / (R_4 + R_6)] q_2 + [R_4 / (R_4 + R_6)] V_0 \quad (2.6b)$$

Now suppose that the R_4 and R_6 are nonlinear. Then we may have difficulty solving the auxiliary equations to get an explicit state form. In general explicit analytic solutions of nonlinear coupled equations are difficult, if not impossible, to achieve.

From the development above, we see that the process of causality assignment is an aid in the process of identifying the algebraic loops in dynamic systems. In Figure 2-1, causality on bonds 4, 5, and 6 can not be determined after assigning causality to source and energy storage elements. Furthermore, algebraic loops in the mathematical sense are

physically related to the existence of dissipation fields. In Figure 2-1, this dissipative field is consisting of nodes R_a , R_b , and associated junction structure. Such fields are called implicit R-fields (IRFs). Reading the partial causality-assigned bond graph of Figure 2-1b, we can easily identify the R fields from other dynamic fields. It is also clear that nonlinear algebraic loops in system equations may prevent the subsequent reduction of the equations to an explicit state-space form. They may make system simulation very difficult to accomplish.

The solution of algebraic loops may not be very important from a theoretical point of view. However, especially for nonlinear systems where an analytical solution is not always possible, solving the loops is very computer-time consuming. Often when the model involves implicit algebraic equations, the approach is to use iterative solution methods. These methods can be very costly and they may introduce difficulties related to the existence and uniqueness of solutions.

Several researchers have been working in a bond graph environment, using three different approaches to these problems. Barreto and Lefevre (1985) try to avoid implicit algebraic loops by modifying the system model. Their basic idea is to consider the implicit part of the model before attempting the simulation and to modify the model suitably so it can be simulated using only explicit methods. Their proposed ways to modify the model include: (1) imposing restrictions in the set of values of admissible solution; (2) changing the model to fit reality.

The second direction is to improve the computational efficiency within the implicit solution framework. Lorenz and Wolper (1985) made an observation on the causality assignment in the case of algebraic loops.

The algebraic loops are found by a related signal graph and the minimum number of independent variables of the algebraic loops is the minimum number of nodes required to break the topological loops. Two general rules are suggested for assigning causality to the implicit R-fields.

The third direction in which the parasitic physical elements are added to eliminate the nonlinear coupled R-field has been employed by Zhou (1984). The method based on bond graph augmentation converts an IRF into a dynamic subsystem that exhibits the proper static characteristics at steady-state and employs a two-time-scale integration technique. This method has a philosophy similar to that of the charge-up method in the ASTAP program (Zeid, 1985). The charge-up method has proven to be very reliable, but it is computationally costly. Since all capacitors and inductors introduced into the original system are assigned a value of 1, for a typical nonlinear circuit, it would take close to $200N$ passes to arrive at a solution, where N is the number of dynamic elements introduced. In Zhou's work, an augmentation sequence and a general rule for parameter selection for arbitrary n -th order subsystems have been suggested. These have been numerically tested in several cases. The algorithm appears promising, but it has not been optimized. Granda (1984) also proposed several approaches to algebraic loops, including adding a storage element.

In this study we assume that a system bond graph containing one or more implicit R fields is given. The task is to solve the coupled algebraic equations efficiently during the system simulation. The decomposition technique is applied to the IRFs and the relevant associated vectors. A direct bond-graph-based algorithm for identifying the minimum number of iteration variables in certain implicit R-field

problem has been developed (Hood, et al., 1987). An advantage of the algorithm is that it works directly with the bond graph itself. This algorithm is also extended to IRFs containing multiport R elements. For an IRF with a general junction structure, a preliminary result on the basis order is developed which facilitates the efficient solution for simulation of system containing such IRFs.

2.2 General Solution by Iterative Methods

A bond graph model may be characterized by the diagram in Figure 2-2 (Rosenberg and Karnopp, 1985). The system inputs are defined by the collection of S_e and S_f elements and are referred to as the source field. Dissipative effects from the R elements of bond graph are grouped into the dissipation field. Dynamic effects are the result of C and I elements in the model and are represented in the energy storage field. Each of these fields has constitutive equations associated with it. They are coupled by a power-conserving connective multiport represented by the junction structure. The Paynter junction structure consists of 0 and 1 junction elements, (It is named after H. M. Paynter, the inventor of the bond graph.) and it is invariant. The transducer field is the collection of the transformers and gyrators, which may have varying moduli. The key vectors of each field are identified in Figure 2-3, where U is the input vector, a function of time; X is the energy variable vector and Z is the coenergy variable vector, the subscripts i and d denote the independent and dependent parts of the storage field; the output vector of the dissipation field D_o is the function of the input vector to the field D_i ; T_i and T_o are the input and output vectors of the modulated junction field.

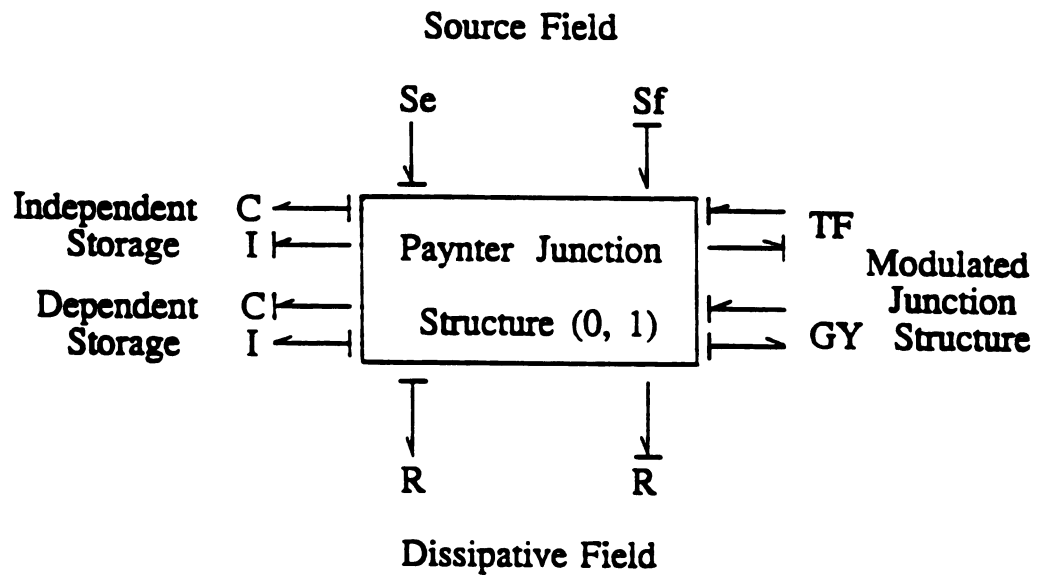


Figure 2-2 Multiport Field Structure

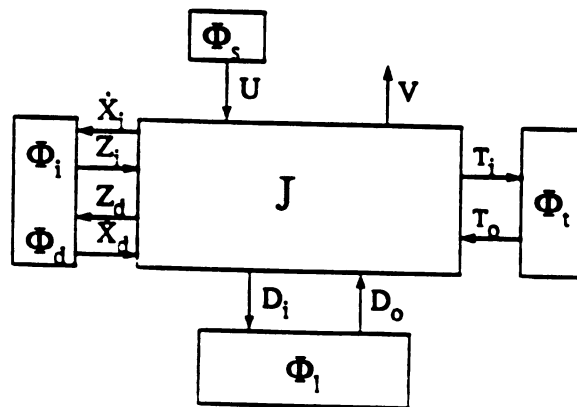


Figure 2-3 Key vectors

When causality has been assigned to the bond graph according to the Sequential Causality Assignment Procedure (SCAP) (Rosenberg and Karnopp, 1983), it is possible to identify each separate implicit R-field within a graph. As a result of assigning causality, a computing diagram can be generated, based on a set of key vectors for various fields in the graph. Ignoring the Z_d , X_d vectors, the general equation structure is given by

$$U = \Phi_u(t) \quad (2.7)$$

$$Z = \Phi_x(X) \quad (2.8)$$

$$D_o = \Phi_L(D_i) \quad (2.9)$$

$$dX/dt = S_{11}Z + S_{13}D_o + S_{14}U \quad (2.10)$$

$$D_i = S_{31}Z + S_{33}D_o + S_{34}U \quad (2.11)$$

for systems with no Z_d , X_d vectors (dependent storages). If implicit R-fields exist in the bond graph, then the matrix S_{33} is non-zero. If S_{33} is zero, then no implicit R-fields exist, and a straightforward procedure for integrating the system equations can be employed. If implicit R-fields exist in the bond graph, then the matrix S_{33} would be non-zero. These fields can be identified and isolated by proper use of causality data. Referring to the field structuring of system equations, the particular subset of equations with which we are concerned is

$$D_o = \Phi_L(D_i) \quad (2.9)$$

$$D_i = S_{31}Z + S_{33}D_o + S_{34}U \quad (2.11)$$

We seek an efficient solution to these equations at each time step, given values for the Z and U vectors. A more succinct form for the equations at a given time is

$$D_o = \Phi_L(D_i) \quad (2.9)$$

$$D_i = S_{33} D_o + C \quad (2.12)$$

or

$$D_i = S_{33}\Phi_L(D_i) + C \quad (2.13)$$

which lead to the single implicit vector equation, where C is a constant vector. Clearly, it would be possible to iterate on D_i to obtain a solution to the problem. There are two useful ways to reach the solution.

(1). Single-partitioned iteration vector

If an R element with port j has its causality assigned uniquely, then the corresponding row in the matrix S_{33} contains all zeros. The value of D_i^j is determined by $D_i^j = C^j$. If an R element belongs to a implicit R -field then its corresponding row in the matrix S_{33} contains nonzero terms. The D_i and D_o vectors are reorganized by gathering all implicit input and output variables into the subvectors D_i^I and D_o^I , respectively, and all explicit input and output variables into the subvectors D_i^E and D_o^E , such that $D_i = [D_i^E, D_i^I]^T$, and $D_o = [D_o^E, D_o^I]^T$. The superscripts E and I denote the explicit and implicit field, respectively. Rearranging the matrix S_{33} , the upper part of the matrix S_{33} is

$$\begin{bmatrix} D_i^E \\ D_i^I \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & S_{33}^I \end{bmatrix} \begin{bmatrix} D_o^E \\ D_o^I \end{bmatrix} + \begin{bmatrix} C^E \\ C^I \end{bmatrix} \quad (2.14)$$

If the dimension of the D_i^E and D_o^E vectors is d_E , the dimension of the D_i^I and D_o^I vectors is d_I , then the dimension of the implicit algebraic equation set to be solved is d_I . That is,

$$D_o^I = \Phi(D_i^I) \quad (2.15a)$$

$$D_i^I = S_{33}^I D_o^I + C^I \quad (2.15b)$$

(3). Multiple-partitioned iteration vector

If the bond graph model contains several implicit R-fields and if the vectors D_i^I and D_o^I are partitioned suitably, the submatrix S_{33}^I is a block diagonal matrix. Equation (2.15b) may be rewritten as

$$\begin{aligned} D_i^{I1} &= S_{33}^{I1} D_o^{I1} + C^{I1} \\ &\cdot \\ &\cdot \\ &\cdot \\ D_i^{Ik} &= S_{33}^{Ik} D_o^{Ik} + C^{Ik} \end{aligned} \quad (2.16)$$

where each subvector has dimension equal to m_i , and $m_1 + m_2 + \dots + m_k = d_I$. The D_i and D_o vectors are sorted into explicit and implicit subsets; then the D_i^I is partitioned according to its fields. This result in the nonzero submatrix of S_{33} having a block diagonal form. Now our task is to solve the following equation set for each local IRF:

$$D_o = \Phi_L(D_i) \quad (2.17)$$

$$D_i = S_{33} D_o + C \quad (2.18)$$

Note that all vectors and matrices in Equations (2.17), (2.18) and the following are local representations.

The advantage of treating the implicit R-fields separately is that the total computation load is smaller. Since the computation cost of an iterative solution to the IRFs appears to vary approximately as the square of the size of the iteration vector, the savings by partitioning into several IRFs can be substantial. Basically, we have a cost of $\alpha(\sum_{i=1}^k m_i^2)$ rather than $\alpha(\sum_{i=1}^k m_i)^2$ by partitioning effectively.

An example is given in Figure 2-4. Two implicit R-fields each contains two 1-port R nodes. The causalities associated with the other four 1-port R nodes are determined uniquely by the SE, C, and I elements. The dimension of the system vectors D_i and D_o is 8. If the D_i , D_o vectors are not partitioned at all, the computation time is 11.43 CPU seconds for a simulation. If the explicit part is separated from the implicit part, then the dimension of iteration vector D_i^I becomes four, and the solution time is reduced to 6.43 CPU seconds. If the D_i^I vector is partitioned into two local IRF vectors, then computation time is reduced to 5.65 CPU seconds. This represents a 50.6 % savings in simulation CPU time by partitioning.

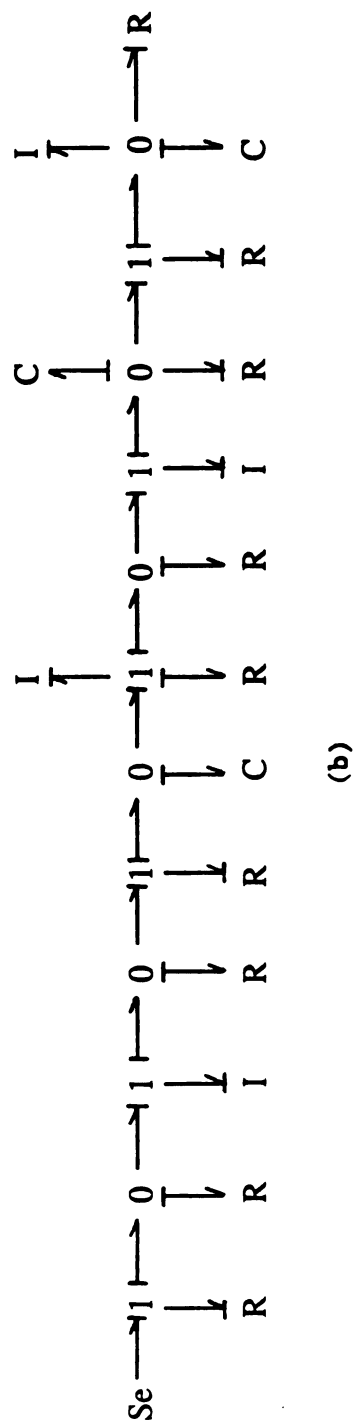
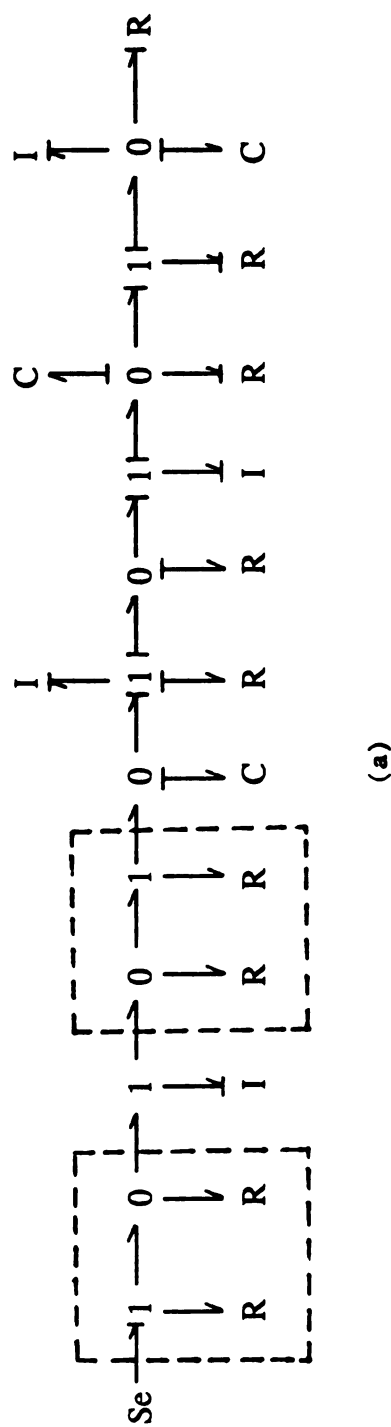


Figure 2-4 A system containing two IRFs

(a) The identified IRFs.

(b) Causality of the IRFs.

2.3 IRFs with 1-port R Elements

An additional refinement for increasing solution efficiency has been developed by Hood, Rosenberg, Withers and Zhou (1987). The method uses earlier work on the basis order for bond graph junction structures (Rosenberg and Moultrie, 1980). In this section we impose two restrictions on the R-field problem, which still leaves us with the most common practical subclass of the general problem. (1) All R nodes in implicit fields are 1-port. (2) The bond graph does not contain transformers (TF) and gyrators (GY) in the implicit fields. The resulting problem is practically important, since many dynamic models of engineering systems are contained in this subclass. Extension to more general problems is discussed subsequently.

Now we ask these questions about the solution of a given IRF. Subject to the restriction imposed above, each obeys local equations of the form of Equations (2.17) and (2.18). We ask "What is the smallest number of iteration variables?" "How can such a set be found?" "How should Equations (2.17) and (2.18) be used for best computing efficiency?"

To answer these questions one of the important properties of the junction structure, basis order, will be used. The basis order is given by two critical numbers, the number of effort inputs (E) and the number of flow inputs (F) required at the junction structure (JS) ports in order to determine all internal effort and flow variables and outputs. The algorithm developed in this section will further reduce the number of iteration variables of an IRF to the minimum of its E and F. This algorithm applies to IRFs that contain only 1-port R elements.

The bonds of a given IRF can be sorted into one of three mutually exclusive sets:

1. the boundary bonds, connecting the IRF to the rest of the graph (these bonds are causal);
2. the R bonds, incident to R nodes, with which are associated the local D_i and D_o vectors; and
3. the remaining bonds, which are internal to the local junction structure (JS).

First we focus on the simple junction structure (SJS) which is composed only of 0- and 1-junction nodes. Earlier work by Rosenberg and Moultrie (1980) has shown that there are two critical numbers associated with a JS. These represent the number of effort (E) and flow (F) inputs required at the JS ports in order to determine all internal variables and the outputs. For completeness we state the rule here for calculating the numbers:

$$E = N_B + N_0 - N_1 - B_0 \quad (2.19a)$$

$$F = N_B + N_1 - N_0 - B_1 \quad (2.19b)$$

where N_B is the number of bonds of the JS;

N_0 is the number of 0-junctions;

N_1 is the number of 1-junctions;

B_0 is the number of bonds incident to the 0-junctions;

B_1 is the number of bonds incident to the 1-junctions.

Next, we observe that a SJS has a pair of separate but related transformations associated with it. Namely, input efforts determine output efforts, and input flows determine output flows. Furthermore, if

all powers at the SJS ports are oriented outward (or all in), the associated matrix S_{33} is skew-symmetric, subject to proper ordering of the port variables (Ort and Martens, 1973, Perelson, 1975). Now assume that causality assignment to the IRF has been completed and is consistent. The proceeding observation allows us to organize Equations (2.17) and (2.18) in detail as follows.

Sort the D_i and D_o vectors into a resistance set (r) and a conductance set (g). The r set has flow inputs to the R ports and effort outputs whose dimensions are E ; the g set has effort inputs to the R ports and flow outputs whose dimensions are F . Write Equation (2.17) as

$$e_r = \Phi_r(f_r) \quad (2.20a)$$

$$f_g = \Phi_g(e_g) \quad (2.20b)$$

where e_r and f_r are associated with the r bond set, and f_g and e_g are associated with the g bond set. Write Equation (2.18) as

$$f_r = S_{rr} e_r + S_{rg} f_g + C_r \quad (2.21a)$$

$$e_g = S_{gr} e_r + S_{gg} f_g + C_g \quad (2.21b)$$

Since the SJS transforms efforts to efforts and flows to flows, then S_{rr} and S_{gg} must be zero. Consequently, we have

$$f_r = S_{rg} f_g + C_r \quad (2.22a)$$

$$e_g = S_{gr} e_r + C_g \quad (2.22b)$$

Furthermore, we note that the combined set (f_r, e_g) contains the SJS outputs, while the combined set (e_r, f_g) contains the SJS inputs. A

computational algorithm may be derived from this observation.

Algorithm

1. Assign causality to the source and storage nodes, using the SCAP.
2. Identify each implicit R-field within the partially causal bond graph.
3. For each implicit R-field:
 - a. Calculate E and F.
 If either E or F is less than one, stop. (There is no guarantee that there are unique outputs from the inputs for the SJS.)
 - b. Obtain a complete, consistent causal orientation for the IRF. (It will obey the E, F numbers.)
 - c. Order the R bonds by resistance (r), then conductance (g) causality. Define the vectors e_r , f_r , f_g , and e_g .
 - d. Assume that E is less than or equal to F. Use e_r as the iteration vector. Make an initial guess e_{ri} for e_r .
 - . Use Equation (2.22b) to find e_g .
 - . Use Equation (2.20b) to find f_g .
 - . Use Equation (2.22a) to find f_r .
 - . Use Equation (2.20a) to find e_r .
 Compare e_{ri} to e_r . If the error is within tolerance, stop. Else return to Equation (2.22b) and repeat the sequence with an updated guess for e_r .
 - . Note: If E is greater than F, use f_g as the iteration vector. The equation order is then (2.22a), (2.20a), (2.22b), (2.20b).

It is observed that the minimum iteration set has the size of $\min(E, F)$. This is always less than or equal to one half the number of R bonds. Reducing the dimension of the iteration vector has a significant positive influence on computing efficiency, as noted previously.

The restrictions placed on the problem structure can be relaxed to a certain extent without changing the algorithm as stated above. A given IRF can contain R nodes with more than one port, provided each such R node is a pure r, or a pure g, type. See Equation (2.20a) and (2.20b) in this regard. In addition, the R-field junction structure can contain transformers (TFs), since they do not alter the structure of the effort-to-effort, flow-to-flow transformation properties. Their effects are combined into S. See Equations (2.22a) and (2.22b). The formulas for E and F need to be modified by introducing the term N_T , which is the number of transformers in the junction structure. The modified formulas are:

$$E = N_B + N_0 - N_1 - B_0 - N_T \quad (2.23a)$$

$$F = N_B + N_1 - N_0 - B_1 - N_T \quad (2.23b)$$

An example of an implicit R-field is shown in Figure 2-5a. The inputs to the IRF are e_a , e_b , and e_c . The goal is to calculate all R-field variables. We first find E and F from the data given in Figure:

$$E = 6 + 1 - 3 - 3 = 1 \quad (2.24a)$$

$$F = 6 + 3 - 1 - 6 = 2 \quad (2.24b)$$

A solution does exist, since both E and F are greater than zero. We obtain a complete, consistent causal orientation, as shown in Figure 2-5b. The ordered r bond vectors are defined, based on the causality, as indicated in the figure. The equations can be written as

$$e_1 = \phi_1(f_1), \quad (2.25)$$

$$f_2 = \phi_2(e_2), \quad (2.26a)$$

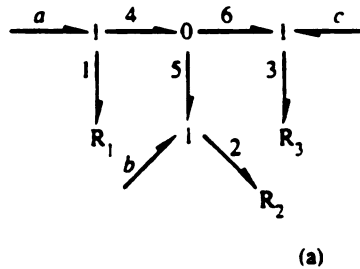
$$f_3 = \phi_3(e_3), \quad (2.26b)$$

$$f_1 = f_2 + f_3, \quad (2.27)$$

$$e_2 = -e_1 + (e_a + e_b), \quad (2.28a)$$

$$e_3 = -e_1 + (e_a + e_c). \quad (2.28b)$$

Since E is less than F, use $e_r = e_1$ as the iteration variable. The equation sequence is (2.28), (2.26), (2.27), (2.25).

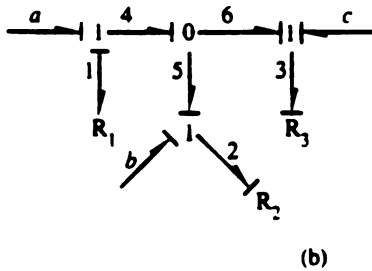


External bonds: a, b, c

R-field bonds: 1 - 6

$N_B = 6, N_0 = 1, N_1 = 3$

$B_0 = 3, B_1 = 6$



$$e_r = [e_1] \quad f_r = [f_1]$$

$$f_g = [f_2, f_3]^t \quad e_g = [e_2, e_3]^t$$

Figure 2-5 An example of an IRF

(a) Bond graph of the R-field

(b) Causality assignment

2.4 IRFs with Multiport R Elements

2.4.1 Multiport Resistances in Physical systems

Multiport resistances arise in many kinds of physical systems, such as hydraulic, thermal, and electronic. A well-known example of multiport resistance can be found in a hydraulic system shown in Figure 2-6a (Rosenberg and Karnopp, 1983). The four-way control valve is characterized by four resistances that are formed by the four edges of the spool lands and corresponding lands in the valve body and modulated by spool position z . The bond graph model is depicted in Figure 2-6b. For a closed-center valve, one can work out the relationships between p_m and Q_s with z as a parameter. These relationships are nonlinear since the pressure drops at individual ports (resistors) are proportional to flow squared. The algebraic reduction of the R-field can be represented in a simplified way (Figure 2-7b) and the constitutive laws for the modulated 2-port R are those shown in Figure 2-7a.

Another example of multiport R element may be found in an electronic circuit. For a 3-terminal element such as the grounded-emitter transistor of Figure 2-8a (Rosenberg and Karnopp, 1983), if we consider the collector-emitter voltage, e_{CE} , base-emitter voltage, e_{BE} , and the currents i_C , i_B , and i_E , the 2-port R representation of Figure 2-8b with a mixed causality has the following constitutive laws:

$$i_C = \phi_1(e_{CE}, i_B) \quad (2.29)$$

$$e_{BE} = \phi_2(e_{CE}, i_B) \quad (2.30)$$

The figure shows a possible connection for a power transistor in a

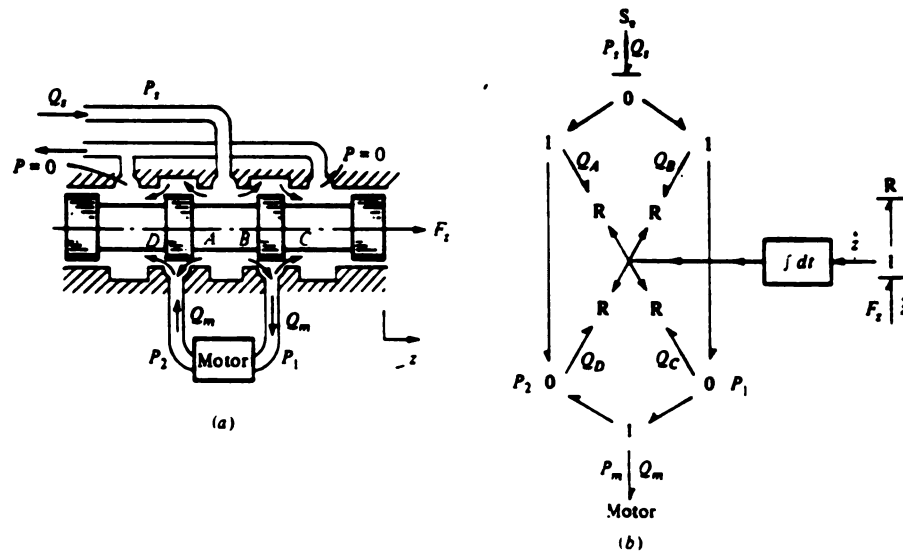


Figure 2-6 Four-way-valve amplifier

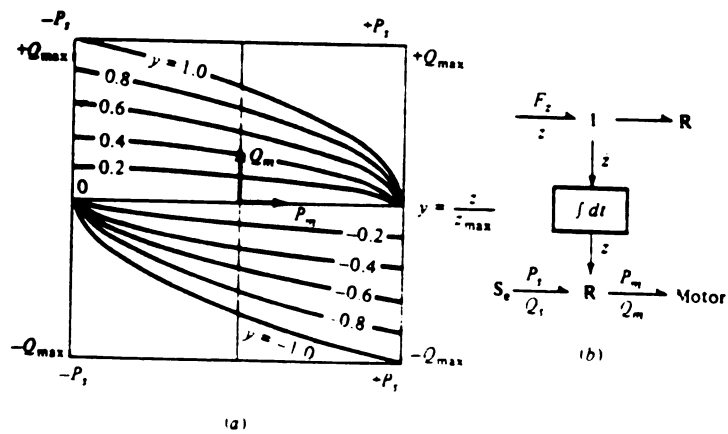


Figure 2-7 Reduced form of hydraulic amplifier

switching mode. The control current i_B is switched rapidly between a (low) blocking value and a (high) conducting value. During the blocking phase, i_C is nearly zero even when e_{CE} takes on quite large values. During the conducting phase e_{CE} is small even for fairly large currents, so that the voltage source is effectively applied directly to the load, the averaged load current can be continuously varied.

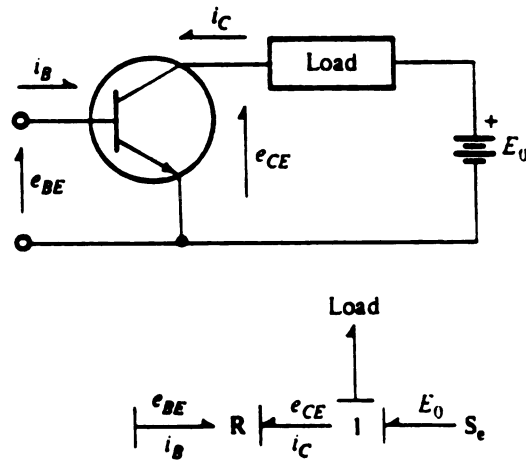


Figure 2-8 Power transistor and its bond graph model

2.4.2 Extension of the Algorithm to an IRF with Multiport R elements

Assume a causal IRF with a number of 1-port R nodes and M multiport R nodes is given as shown in Figure 2-9, where R denotes a set of R nodes in resistant form (R), G denotes a set of R nodes in conductance form (G), while RG_i denotes a multiport R node with mixed causality. In addition, U denotes the input variables, while V denotes the output variables. S characterizes the junction structure. The key vectors are:

$$\begin{aligned} f_r &= [f_{ro}, f_{r1}, \dots, f_{rM}]^t; \\ e_g &= [e_{go}, e_{g1}, \dots, e_{gM}]^t; \\ e_r &= [e_{ro}, e_{r1}, \dots, e_{rM}]^t; \\ f_g &= [f_{go}, f_{g1}, \dots, f_{gM}]^t; \end{aligned}$$

and in an aggregated form,

$$\begin{aligned} D_i &= [f_r, e_g]^t; \\ D_o &= [e_r, f_g]^t \end{aligned}$$

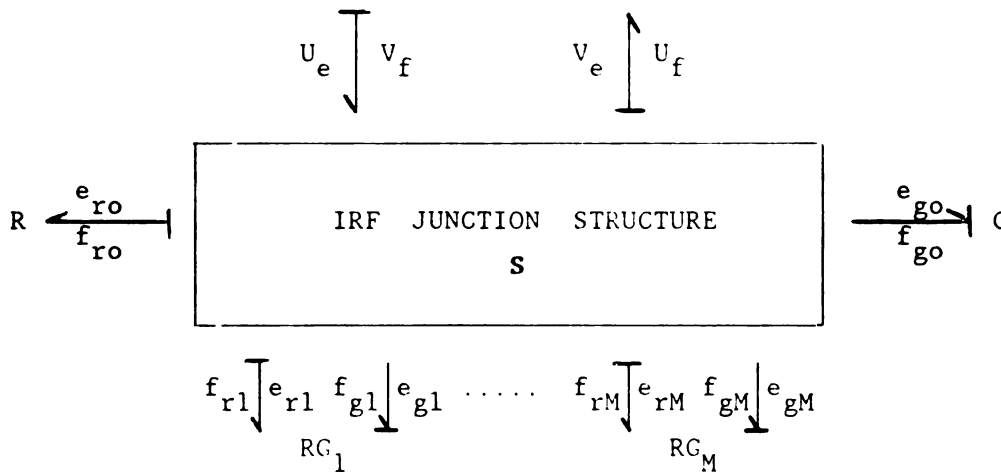


Figure 2-9 Key vectors in an IRF

where f_r is the flow input vector to the R ports; e_g is the effort input vector to the G ports; e_r is the effort output vector from the R ports; and f_g is the flow output vector from G ports. D_i is the input vector to R nodes and D_o is the output vector from R nodes. The vectors with subscript o are associated with the 1-port R nodes. The vectors with subscript m are associated with the m-th multiport R node, where $m = 1, 2, \dots, M$. The constitutive equations for the 1-port R nodes are:

$$e_{ro} = \Phi_{ro}(f_{ro}) \quad (2.31a)$$

$$f_{go} = \Phi_{go}(e_{go}) \quad (2.31b)$$

and the constitutive equations for multiport R nodes are:

$$e_{rm} = \Phi_{rm}(f_{rm}, e_{gm}) \quad (2.32a)$$

$$f_{gm} = \Phi_{gm}(f_{rm}, e_{gm}) \quad m = 1, 2, \dots, M \quad (2.32b)$$

The input D_i and the output D_o are linearly related through the junction structure of the IRF. The relations between D_i and D_o are represented by the junction structure matrix S for the IRF, namely,

$$D_i = S D_o + C \quad (2.33)$$

where C is a constant vector derived from input vector U . Assuming that the powers on all bonds of the R-field are directed towards the R nodes, the local junction structure matrix S has the following form:

$$S = \begin{vmatrix} 0 & S_f \\ S_e & 0 \end{vmatrix} = \begin{vmatrix} 0 & S_{rogo} & S_{rogl} & \dots & S_{rogM} \\ S_{rlgo} & S_{gorgl} & \dots & S_{gorM} & \\ \vdots & \vdots & & \vdots & \\ S_{rMgo} & \dots & \dots & S_{rMgM} & \\ S_{gMro} & \dots & \dots & S_{gMrM} & 0 \end{vmatrix}$$

where S is skew-symmetric.

The connective Equation (2.33) may be written in some detail as:

$$f_r = S_f f_g + C_r \quad (2.34a)$$

$$e_g = S_g e_r + C_g \quad (2.34b)$$

We may also simply rewrite the equations in a linear function form by:

$$f_{ro} = L_{ro}(f_{go}, f_{gl}, \dots, f_{gM}) + C_{ro} \quad (2.35a)$$

$$f_{rm} = L_{rm}(f_{go}, f_{gl}, \dots, f_{gM}) + C_{rm} \quad (2.35b)$$

and

$$e_{go} = L_{go}(e_{ro}, e_{rl}, \dots, e_{rM}) + C_{go} \quad (4.36a)$$

$$e_{gm} = L_{gm}(e_{ro}, e_{rl}, \dots, e_{rM}) + C_{gm} \quad (2.36b)$$

where $m = 1, 2, \dots, M$.

We now seek as an iteration set a minimum satisfactory set of variables. If we construct a functional diagram to represent the mathematical relations among the variables based on Equations (2.31), (2.32), (2.35), and (2.36), it appears as in Figure 2-10. The weight of

each edge is the dimension of the associated vector. From this diagram, the algebraic loops can be identified and a minimum iteration set may be determined by removing a set of edges which break all loops and whose sum of weights is a minimum. However, for our purpose of identifying the minimum iteration variables, a simpler variable flow diagram in Figure 2-11 can be generated such that a vertex represents a vector and an edge represents a condition of input-output. This diagram emphasizes the computation requirement of which input vectors are needed to generate a particular output vector without any detail of their functional relations. For convenience in our initial development let us ignore the existence of e_o , f_o temporarily.

The algebraic loops can be identified from this variable flow diagram. The iteration variable set with minimum dimension can be found from its associated loops in the variable flow diagram. The removal of the vertices (iteration variable set) from the variable flow diagram destroys their associated loops. If the selection can be guided by some rules, it will facilitate the automatic computation efficiently. For consistency, we always select the iteration variables from the input variables to the junction structure, namely, e_r and f_g .

Every set consisting of four vertices representing the vectors (e_r , f_r , e_g , and f_g)_m and the edges among them is defined as a block of the digraph. Each block connects (M-1) other blocks with the same structure by $2*(M-1)$ outgoing edges departing from vertices f_g and e_r , while the vertices e_g and f_r connect the (M-1) blocks by $2*(M-1)$ incoming edges.

All blocks are arranged as shown in Figure 2-11 such that all the vertices (vectors) associated with r-type causality are located at the

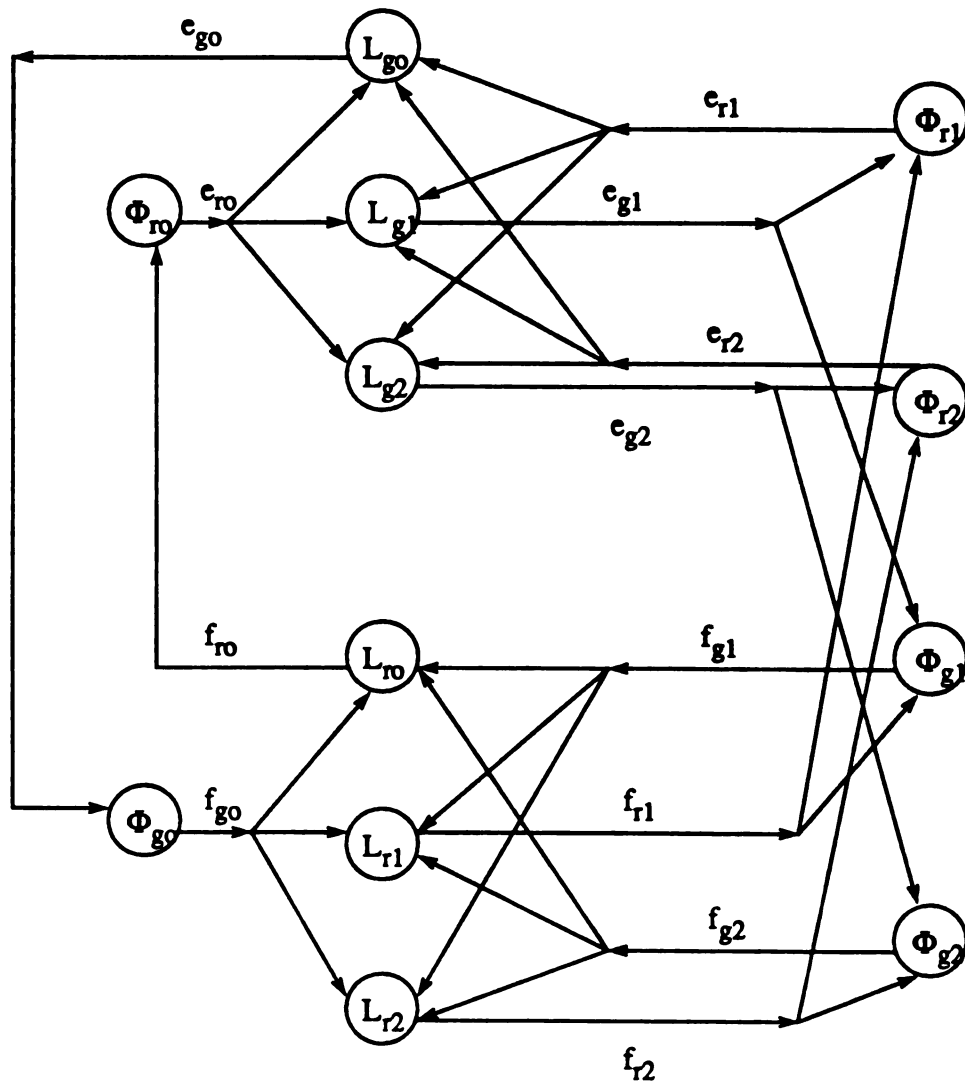


Figure 2-10 Functional diagram of an IRF

right side, while all the vertices (vectors) associated with g-type causality are located on the left side of the digraph. The vertex associated with r-type causality is classified as r-type vertex, while the vertex associated with g-type causality is classified as g-type vertex. In order to break all loops in the digraph, it is obvious that one of the two vertices in each row (either $(e_g, f_g)_m$ or $(e_r, f_r)_m$) must be removed since it is necessary to destroy the loops formed by these two-vertex sets.

Referring to this digraph with a repeated and symmetrical pattern, every loop in this digraph contain both g-type and r-type vertices. Therefore, removal of either all the r-type or all the g-type vertices breaks all loops. Thus (e_g, f_g) or (e_r, f_r) can be chosen as a feasible solution. The size of the selected vertices (vector) set is two times the dimension of (e_g, f_g) or (e_r, f_r) . i.e., $2F$ or $2E$, respectively. Considering the number of the basis order, E and F , the set of iteration variables with smaller size can be chosen. Thus the size is given by

$$N = 2 * \min \{ E, F \} \quad (2.37)$$

For convenience in the next development, let us assume the case of $E \leq F$, in which all the vertices on the right side are removed (If $F \leq E$, then certain arguments are modified suitably.). Although $E \leq F$, this does not guarantee that in every block m , $N_r \leq N_g$ is true. Since the e_{rm} s are the inputs of the junction structure, let us keep these in our solution temporarily.

Under this condition for a particular block with e_r removed, removing f_g instead of f_r also ensures no loops associated with this

particular block, because there are no more outgoing edges from this block. A size reduction may be achieved by removal of one of the two vertices $(f_g, f_r)_m$ with smaller size. By applying this step to each block, N can be reduced from Equation (2.37) to

$$N = \min \{ E, F \} + \sum_{m=1}^M \min \{ N_r, N_g \}_m \quad (2.38)$$

Can we reduce N further? Suppose that there exists an i -th block such that $N_{gi} < N_{ri}$. If vertex e_{gi} is removed instead of e_{ri} , at least one loop will be found in the remaining digraph (Figure 2-12). The reason is the following. Since $N_{gi} < N_{ri}$, the vertex f_{gi} has been removed, while f_{ri} is kept in the last step. However, since $E \leq F$, we always can find at least one block, say the j -th block, such that $N_{rj} \leq N_{gj}$. In this block, vertices e_{rj} and f_{rj} have been removed and vertices e_{gj} and f_{gj} remain. It is easy to verify that a loop $(e_{ri}, e_{gj}, f_{gj}, f_{ri}, e_{ri})$ exists in the remaining digraph. This shows that such exchange of vertices e_{gi} and e_{ri} is not allowed, even though $N_{gi} \leq N_{ri}$. Thus no further reduction on N can be achieved.

Now consider the existence of 1-port R nodes. These are represented by the block consisting of the vertices e_{ro} , f_{ro} , e_{go} , and f_{ro} as shown in Figure 2-13. It is a special case of the general blocks, namely, no two-vertex loops exist. In the case of $E < F$, removing only e_{ro} can break all possible loops associated with this block. Therefore, incorporating this advantage into the result, the smallest size of the iteration variables for an IRF containing both 1-port and multiport R nodes with assigned causalities is given by

$$N = \min \{ E, F \} + \sum_{m=1}^M \min \{ N_r, N_g \}_m \quad (2.39)$$

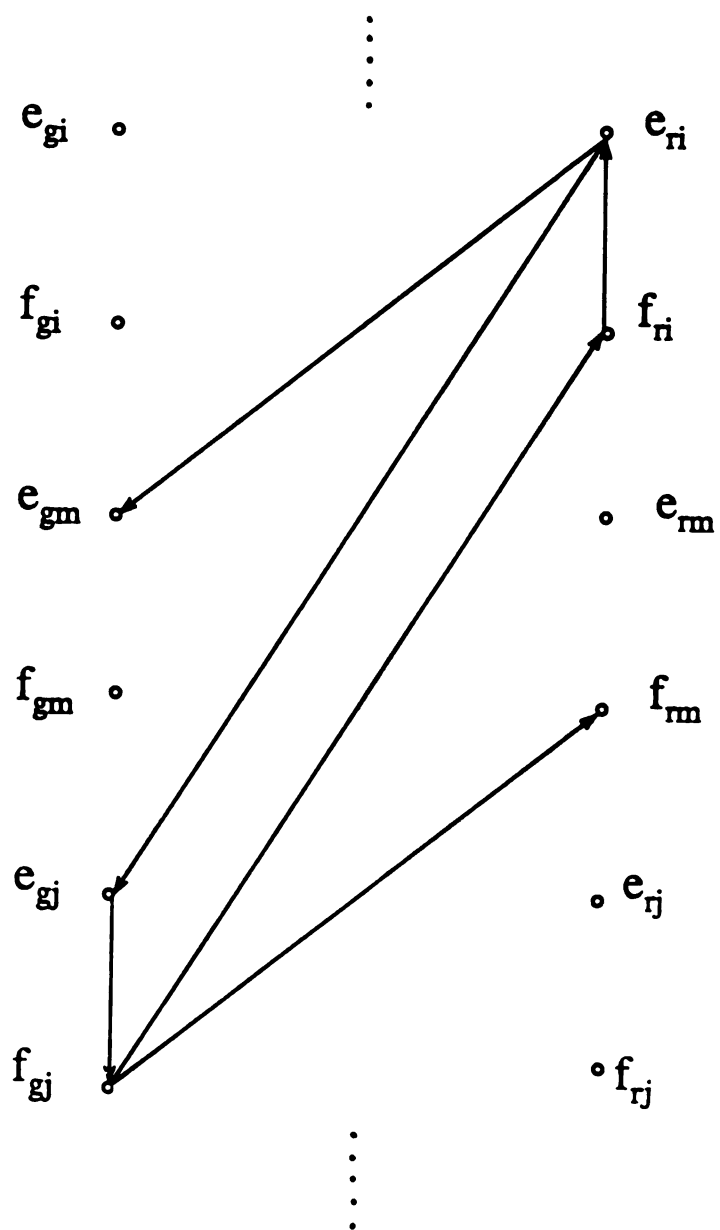


Figure 2-12 Loop formed by an improper vertex exchange

Table 2-1 Possible variable selection

	$F \leq E$	$E \leq F$
	$(\sum_{m=0}^M U_{gm} f_{gm})$	$(\sum_{m=0}^M U_{rm} e_{rm})$
$N_{gm} \leq N_{rm}$	$+ e_{g1}$	$+ f_{g1}$
$N_{gm} \geq N_{rm}$	$+ e_{r1}$	$+ f_{r1}$

Note here that $E = \sum_{m=0}^M N_{rm}$ and $F = \sum_{m=0}^M N_{gm}$. An iteration variable set with this property can be identified.

From the analysis above, for each multiport R node four possibilities exist, which are listed in Table 2-1. A simple rule to determine the minimum iteration variable set for the IRF with multiport R elements is derived from the variable flow diagram and Table 2-1 as follows:

- 1) If $F \leq E$, choose $(\sum_{m=0}^M U_{gm} f_{gm})$. else $(\sum_{m=0}^M U_{rm} e_{rm})$;
- 2) For each multiport R node:
 - a. In addition to $(\sum_{m=0}^M U_{gm} f_{gm})$, if $N_{rm} \leq N_{gm}$, choose e_{rm} ,
else e_{gm} .
 - b. In addition to $(\sum_{m=0}^M U_{rm} e_{rm})$, if $N_{rm} \leq N_{gm}$, choose f_{rm} ,
else f_{rm} .

An example

A simple electronic circuit is shown in Figure 2-14a. After assigning causality to the Sf and Se elements, an IRF is identified (Figure 2-14b). By assigning suitable causality to ports 3 and 5, a causally completed bond graph is obtained (Figure 2-14c).

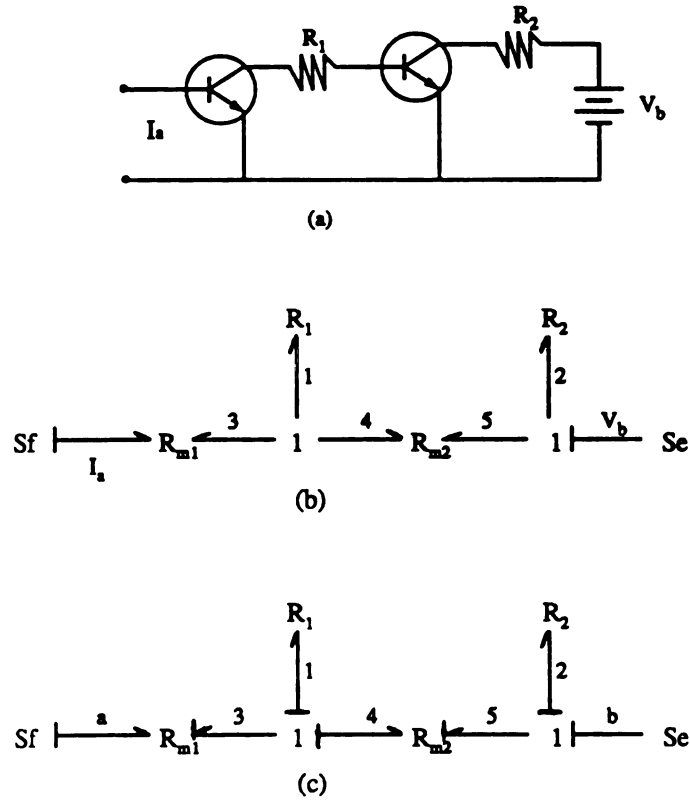


Figure 2-14 An electronic circuit and its bond graph model

The vectors associated with 1-port R nodes are

$$e_{ro} = [e_1, e_2]^t$$

$$f_{ro} = [f_1, f_2]^t$$

and

$$e_{go} = f_{go} = [0]$$

The vectors associated with multiport R nodes are

$$e_{r1} = f_{r1} = [0]$$

$$f_{g1} = [f_3]$$

$$e_{g1} = [e_3]$$

and

$$e_{r2} = [e_4]$$

$$f_{r2} = [f_4]$$

$$f_{g2} = [f_5]$$

$$e_{g2} = [e_5]$$

we may calculate the numbers E and F by

$$E = \sum_{m=0}^2 \dim (e_{rm}) = 2 + 0 + 1 = 3 \quad (2.40a)$$

$$F = \sum_{m=0}^2 \dim (f_{gm}) = 0 + 1 + 1 = 2 \quad (2.40b)$$

The minimum number of sufficient iteration variables is determined by

$$\begin{aligned} N &= \min \{ E, F \} + \sum_{m=1}^M \min \{ N_{rm}, N_{gm} \} \\ &= \min \{ 3, 2 \} + \min \{ 0, 1 \} + \min \{ 1, 1 \} = 3 \quad (2.41) \end{aligned}$$

The iteration variables are chosen by the rule as follows: f_3 , f_5 , and e_5 . The constitutive equations for the R nodes and the junction structure are

$$D_o \quad e_{ro}: \quad e_1 = \phi_1(f_1) \quad (2.42a)$$

$$e_2 = \phi_2(f_2) \quad (2.42b)$$

$$f_{go}: \quad \text{-----}$$

$$e_{r1}: \quad \text{-----}$$

$$f_{g1}: \quad f_3 = \phi_3(f_a, e_3) \quad (2.43)$$

$$e_{r2}: \quad e_4 = \phi_4(f_4, e_5) \quad (2.44)$$

$$f_{g2}: \quad f_5 = \phi_5(f_4, e_5) \quad (2.45)$$

$$D_i \quad f_{ro}: \quad f_1 = f_3 \quad (2.46a)$$

$$f_2 = f_5 \quad (2.46b)$$

$$e_{go}: \quad \text{-----}$$

$$f_{r1}: \quad \text{-----}$$

$$e_{g1}: \quad e_3 = -e_1 - e_4 \quad (2.47)$$

$$f_{r2}: \quad f_4 = f_3 \quad (2.48)$$

$$e_{g2}: \quad e_5 = e_b - e_2 \quad (2.49)$$

The iteration variables are $f_{g0} \cup f_{g1} \cup f_{g2} \cup e_{g2}$, namely, $(f_3, f_5, \text{ and } e_5)$. The iteration process may start with a set of initial guess $(f_{3i}, f_{5i}, \text{ and } e_{5i})$ to (f_3, f_5, e_5) . The iteration sequence order is (2.46), (2.48), (2.42), (2.44), (2.47), (2.49), (2.45), (2.43). A new set of f_3 , f_5 , and e_5 are generated and become updated initial values. Thus the iteration can be repeated until the errors are within specified tolerance.

2.4.3 Assigning causality to an IRF Containing one Multiport R node

From the study of iteration variables for a causal IRF containing a multiport R node, it is clear that if the multiport R node has its causality assigned in such a way that it is in either complete R-form or complete G-form (i.e., either $N_{rm} = 0$ or $N_{gm} = 0$), the number of iteration variables is minimum and equal to the smaller of E and F. If such a causal assignment is not possible, we should make the additional iteration variable set as small as possible. That is, either N_{rm} or N_{gm} is the minimum. How can we determine the minimum number of N_{rm} or N_{gm} ? How do we assign causality so that the minimum number of iteration variables can be used? This subsection addresses these questions.

Figure 2-15 is a diagram for an acausal IRF where $R_1, \dots, R_n$ are 1-port R nodes and the large R is a multiport R node. U_e and U_f are input vectors to the IRF and V_f and V_e are output vectors.

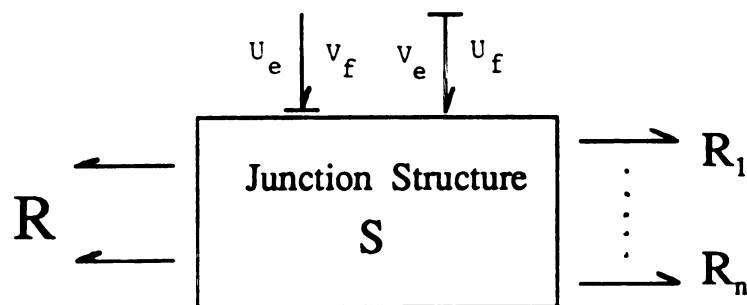


Figure 2-15 Acausal IRF with multiport R node

Suppose a multiport R node has n ports. If we start at one of the ports to trace out a path consisting of alternating bonds and junction nodes (0, 1, TF), the path may end up at a 1-port R node or the multiport R node itself. A path starting at and ending up at the same multiport R node is called a self-loop. Several paths may share some common ports and common junction nodes. The collection of the ports of such coupled self-loops and all the junction nodes in the paths, and all 1-port R nodes adjacent to the junction nodes, is defined as an R-block. Figure 2-16 shows an implicit R-field containing a multiport R node which has a causally determined bond 7 and three acausal bonds 4, 5, and 6. A self-loop is identified, that is $(R_m - 0a - 1a - 0b - 1b - R_m)$. The R-block associated with this self-loop consists of node R_m , $0a$, R_2 , $1a$, $0b$, $1b$, and R_3 .

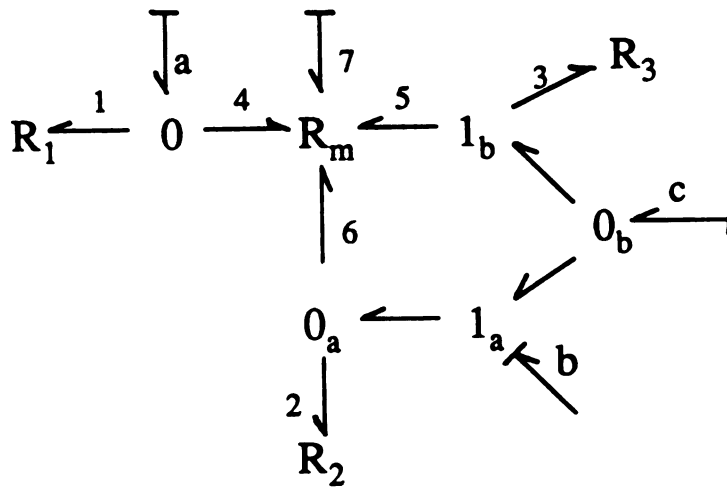


Figure 2-16 An example of self-loop and R-block

Before we make a further discussion, we need to have the following assumptions. First, no redundant R ports exist in the IRF, namely, no two or more 1-port R nodes are allowed to be adjacent to a common junction node, since such 1-port nodes may be aggregated into a single 1-port R node. Second, no short loop exists among the multiple R ports, namely, no two or more ports of a multiport R node are adjacent to a common junction node. This has the same reason as for the first assumption.

It is easy to show that a port of the multiport R node, which is not contained in any R-blocks, can be assigned with either R-form or G-form causality. For example, bond 4 of the multiport R in Figure 2-16 is not in any R-block, it can be assigned with either effort in or effort out. However, the ports in an R-block, for example, bond 5 and 6, may or may not be assigned freely depending on the local junction structure of the R-block.

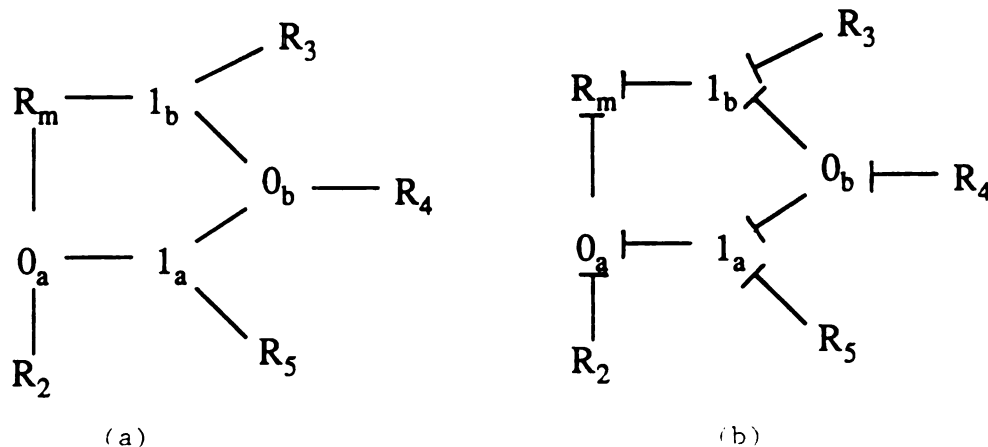
Each R-block has its own local basis order pair. Let M_i denote the number of ports associated with the multiport R node in the i -th R-block. Let E_i and F_i denote the basis order pair for the i -th R-block. Given E and F for a SJS or a WJS, not any subset of ports can always be assigned in, say, R causality. However, if we assume that any subset of ports in an R-block with given E and F can be assigned desired causality obeying E and F , we can derive a possible minimum number of iteration variables and guide the causality assignment from the discussion on four types of R-blocks.

Type 1. M_i is less than or equal to both E_i and F_i .

Since no short loop is allowed in the bond graph, tracing the causal orientation from a port of the multiport R node will always end up at a 1-port R node. The M_i ports of the multiport R node can be assigned either all in R-form or G-form. Figure 2-17 illustrates an R-block of type 1. The local basis order is calculated to be $E = 3$ and $F = 3$. The number of acausal multiport bonds in the R-block is 2. These two ports can be assigned either in R-form or G-form.

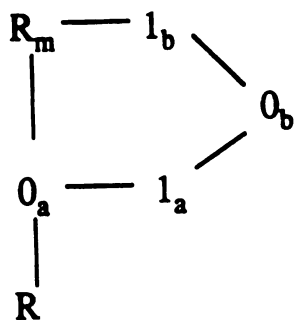
Type 2. M_i is greater than E_i but less than or equal to F_i .

In this case, M_i ports of the multiport R node can be assigned all in G-form, but not in R-form. In Figure 2-18, since $E_1 = 1$, and $F_1 = 2$, the two multiports can not be assigned as all R-form, but it may be assigned as all G-form.



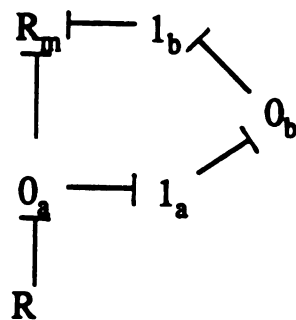
$$\begin{array}{lll}
 N_B = 9, & N_0 = 2, & N_1 = 2, \\
 B_0 = 6, & B_1 = 6, & M_1 = 2, \\
 E_1 = 9 + 2 - 2 - 6 = 3, & & E_1 > M_1 \\
 F_1 = 9 + 2 - 2 - 6 = 3, & & F_1 > M_1
 \end{array}$$

Figure 2-17 An R-block of type 1



(a)

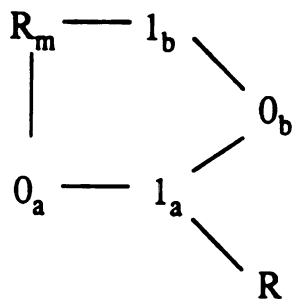
$$\begin{aligned}
 N_B &= 6, & N_0 &= 2, & N_1 &= 2, \\
 B_0 &= 5, & B_1 &= 4, & M_1 &= 2, \\
 E_1 &= 6 + 2 - 2 - 5 = 1 \\
 F_1 &= 6 + 2 - 2 - 4 = 2
 \end{aligned}$$



(b)

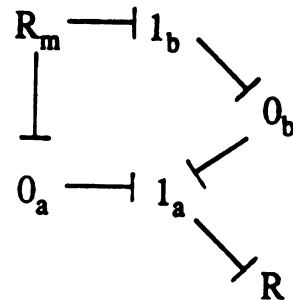
$$\begin{aligned}
 E_1 &< M_1 \\
 F_1 &\geq M_1
 \end{aligned}$$

Figure 2-18 An R-block of type 2



(a)

$$\begin{aligned}
 N_B &= 6, & N_0 &= 2, & N_1 &= 2, \\
 B_0 &= 4, & B_1 &= 5, & M_1 &= 2, \\
 E_1 &= 6 + 2 - 2 - 4 = 2 \\
 F_1 &= 6 + 2 - 2 - 5 = 1
 \end{aligned}$$



(b)

$$\begin{aligned}
 E_1 &\geq M_1 \\
 F_1 &< M_1
 \end{aligned}$$

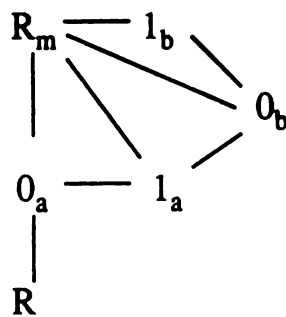
Figure 2-19 An R-block of type 2

Type 3. M_i is greater than F_i but less than or equal to E_i .

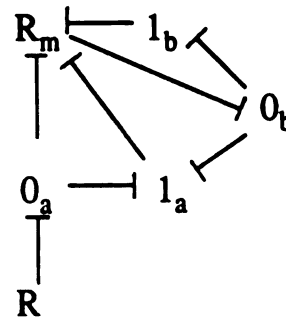
The result is opposite to that in case 2. The R-form can be assigned on all M_i ports, or F_i ports can be in G-form and the remaining ports must be in R-form. The causality assignment can be seen from the example in Figure 2-19.

Type 4. M_i is greater than both E_i and F_i .

Obviously, the M_i cannot be assigned either complete R-form or G-form. They can be assigned E_i R-form causalities and $(M_i - E_i)$ G-form causalities, or inversely, F_i G-form and $(M_i - F_i)$ R-form causalities. It is shown in Figure 2-20.



(a)



(b)

$$N_B = 8, \quad N_0 = 2, \quad N_1 = 2.$$

$$B_0 = 6, \quad B_1 = 5, \quad M_1 = 4.$$

$$E_1 = 8 + 2 - 2 - 6 = 2$$

$$F_1 = 8 + 2 - 2 - 5 = 3$$

$$E_1 < M_1$$

$$F_1 < M_1$$

Figure 2-20 An R-block of type 4

Let N_0 denote the number of free ports of the multiport R node, which are not in any R-blocks. An IRF is decomposed and all R-blocks are identified and classified according to above types. Let E_i , F_i and M_i denote the basis order pair and the number of ports of the multiport R node in the i -th type-1 R-block, respectively, where $i = 1, \dots, I$. Let E_j , F_j and M_j denote the basis order pair and the number of ports of the multiport node in the j -th type-2 R-block, respectively, where $j = 1, \dots, J$. Let E_k , F_k and M_k denote the basis order pair and the number of ports of the multiport node in the k -th type-3 R-block, respectively, where $k = 1, \dots, K$. Let E_l , F_l and M_l denote the basis order pair and the number of ports of the multiport node in the l -th type-4 R-block, respectively, where $l = 1, \dots, L$. For a particular multiport R node, we may try to assign as many R-form causalities as the structure allows. This process is called the maximum R-form assignment. The maximum G-form assignment is just its opposite.

In the maximum R-form assignment, the maximum possible number of effort outputs and the minimum possible number of flow outputs from the multiport R can be determined by

$$E_r = \sum_{i=1}^I N_i + \sum_{j=1}^J E_j + \sum_{k=1}^K M_k + \sum_{l=1}^L E_l + N_0 \quad (2.50a)$$

$$F_r = \sum_{j=1}^J (M_j - E_j) + \sum_{l=1}^L (M_l - E_l) \quad (2.50b)$$

where E_r is the maximum possible number of ports with R causality and F_r is the minimum possible number of ports G causality for the multiport R node.

In the maximum G-form assignment, the minimum possible number of effort outputs and maximum possible flow outputs from the multiport R node can be determined by

$$E_g = \sum_{k=1}^K (M_k - F_k) + \sum_{l=1}^L (M_l - F_l) \quad (2.51a)$$

$$F_g = \sum_{i=1}^I M_i + \sum_{j=1}^J M_j + \sum_{k=1}^K F_k + \sum_{l=1}^L F_l + N_0 \quad (2.51b)$$

where E_g is the minimum possible number of ports with R causality and F_g is the maximum possible number of ports with G causality for the multiport R node.

The smallest dimension set is found from F_r and E_g . If $F_r \leq E_g$, then the maximum R-form assignment would be used and the number of iteration variables of IRF would be the minimum. Since E_g is the minimum number of the R-form bonds, it is the smallest possible size of the vector e_{r1} , i.e. N_{r1} in Rule 1; since F_r is the minimum possible number of the G-form bonds, it is the smallest possible size of f_{g1} , i.e. N_{g1} in Rule 1. By comparing E_g and F_r , we find smallest additional iteration set to the set determined by $\min \{E, F\}$. In general, The minimum possible number of iteration variables is determined by:

$$N = \min \{E, F\} + \min \{E_g, F_r\}_m \quad (2.52)$$

This result may be applied to guide the causality assignment for the multiport R node in computation automation. Let us discuss the implications of the above result.

1. If there are no R-blocks, F_r and E_g are equal to zero. This fact suggests that the causalities of the multiport R node can freely be assigned in either complete R-form or G-form.

2. If only type-1 R-blocks exist, F_r and E_g are also equal to zero. (Same as in case 1.)

3. If neither a type-3 nor a type-4 R-block exists, but type-1 and/or type-2 R-blocks exist, $E_g = 0$ and the multiport R node can have complete G-form causalities.

4. If neither a type-2 nor a type-4 R-block exists, but type-1 and/or type-3 R-blocks exist, then $F_r = 0$ and the multiport R node can have complete R-form causalities.

5. If a type 4 R-block exists, then $F_r \neq 0$ and $E_g \neq 0$. The multiport R node will not be in complete R-form or G-form for sure.

From the discussion above, if there is no self-loop of the multiport R node, the multiport R node is not structurally coupled (Figure 2-21). In this case, from the result above, it is easy to see that the multiport R can always be put into either R-form or G-form completely. The causality can be assigned according to the required physical laws of the multiport R node. The minimum number of sufficient iteration variables can be determined by the previous rule, i.e.

$$N = \min \{ E, F \} \quad (2.53)$$

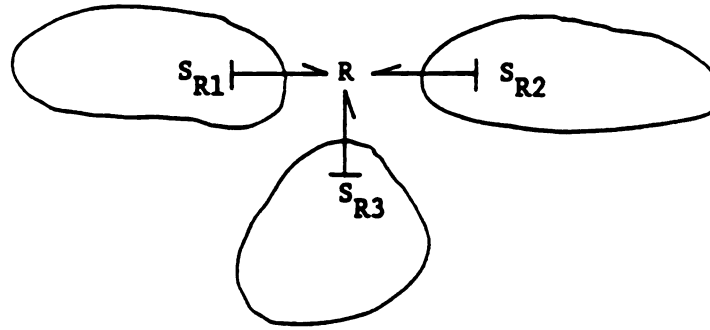


Figure 2-21 A multiport R node without structural coupling

For example, if the multiport R node is in complete R-form, the constitutive and system equations will be reduced to a simpler form

$$e_{ro} = \Phi_{ro}(f_{ro}) \quad (2.54a)$$

$$e_{rl} = \Phi_{rl}(f_{rl}) \quad (2.54b)$$

$$f_{go} = \Phi_{go}(e_{go}) \quad (2.55)$$

$$f_{ro} = S_{rogo} f_{go} \quad (2.56a)$$

$$f_{rl} = S_{rlgo} f_{go} \quad (2.56b)$$

$$e_{go} = S_{goro} e_{ro} + S_{rlgo} e_{rl} \quad (2.57)$$

If the multiport node is associated with self-loops, then this node is structurally coupled. The multiport R node can or cannot be put in either R-form or G-form completely, depending on the structure of the IRF. If the multiport R node can not be put into either R-form or G-form (i.e., mixed form), the coupling among the ports of the multiport R node is called essential structural coupling. If the causality is assigned, we will find a causal loop which starts at one of the multiport node and ends up at another port of the same multiport node.

A strategy for assigning causality to an IRF containing one multiport R node is stated as follows:

1. Assign causality to the source and storage nodes, using the SCAP.
2. Identify each implicit R-field within the partially causal bond graph.
3. For each implicit R-field:
 - a. Calculate E and F.
 If either E or F is less than one, stop. (There is no guarantee that there are unique outputs from the inputs for the IRF.)
 - b. If a multiport R node exists, identify all self-loops and associated R-blocks. For each R-block:
 - . Calculate the local basis order $(E, F)_m$ and M_m .
 - . Classify the type for the R-block.
 - c. Calculate the pair of $(E_g, F_r)_m$ and determine the possible minimum number of iteration variables by

$$N = \min \{E, F\} + \sum_{i=m}^M \min \{E_g, F_r\}_m$$
 - d. Assign causality to multiport R node first according to $\min (E_g, F_r)_m$. Obtain a complete, consistent causal orientation for the IRF. (It will obey the E, F numbers and $(E, F)_n$.)
 - e. Order the R bonds by resistance (r), then conductance (g) causality. Define the vectors e_r , f_r , f_g , and e_g .
 - f. Assume that E is less than or equal to F and E_g is less than or equal to F_r , then $N = E + E_g$. Use e_{r0} , e_{r1} , and f_{r1}

as the iteration vectors. Make an initial guess e_{roi} , e_{rli} , and f_{rli} for e_{ro} , e_{rl} , and f_{rl} .

. Use Equation (2.36) to find e_{go} and e_{gl} .

. Use Equation (2.31b) to find f_{go} .

. Use Equation (2.32a and b) to find e_{rl} and f_{gl} .

. Use Equation (2.35) to find f_{ro} and f_{rl} .

. Use Equation (2.31a) to find e_{ro} .

Compare e_{roi} , e_{rli} , and f_{rli} to e_{ro} , e_{rl} , and f_{rl} . If the error is within tolerance, stop.

Else return to Equation (2.36) and repeat sequence with the updated guess for e_{ro} , e_{rl} , and f_{rl} .

Note:

1. If E is less than F and F_r is less than E_g , use e_{ro} , e_{rl} , and f_{gl} as the iteration vectors. The equation order is then (2.36), (2.31b), (2.35), (2.31a), (2.32).

2. If F is less than E and E_g is less than F_r , use f_{go} , f_{gl} , and e_{rl} as the iteration vectors. The equation order is then (2.35), (2.31a), (2.32), (2.36), (2.31b).

3. If F is less than E and F_r is less than E_g , use f_{go} , f_{gl} , and e_{gl} as the iteration vectors. The equation order is then (2.35), (2.31a), (2.32), (2.36), (2.31 b).

The following two examples show how this strategy works. The first example in Figure 2-22a depicts an isolated IRF with one multiport R node. The pair of basis orders are calculated first:

$$E = 8 + 2 - 3 - 5 = 2 \quad (2.58a)$$

$$F = 8 + 3 - 2 - 6 = 3 \quad (2.58b)$$

At this stage we do not know how to assign causality to this IRF and also we are not sure that the number $E = 2$ is the number of the iteration variables. Now, a self-loop ($R_m - 0a - 1a - 0b - 1b - R_m$) is identified, and so the associated R-block ($R_m, 0a, 1a, 0b, 1b, R_5$). The local basis orders of the R-block is obtained by:

$$E_1 = 6 + 2 - 2 - 5 = 1 \quad (2.59a)$$

$$F_1 = 6 + 2 - 2 - 4 = 2 \quad (2.59b)$$

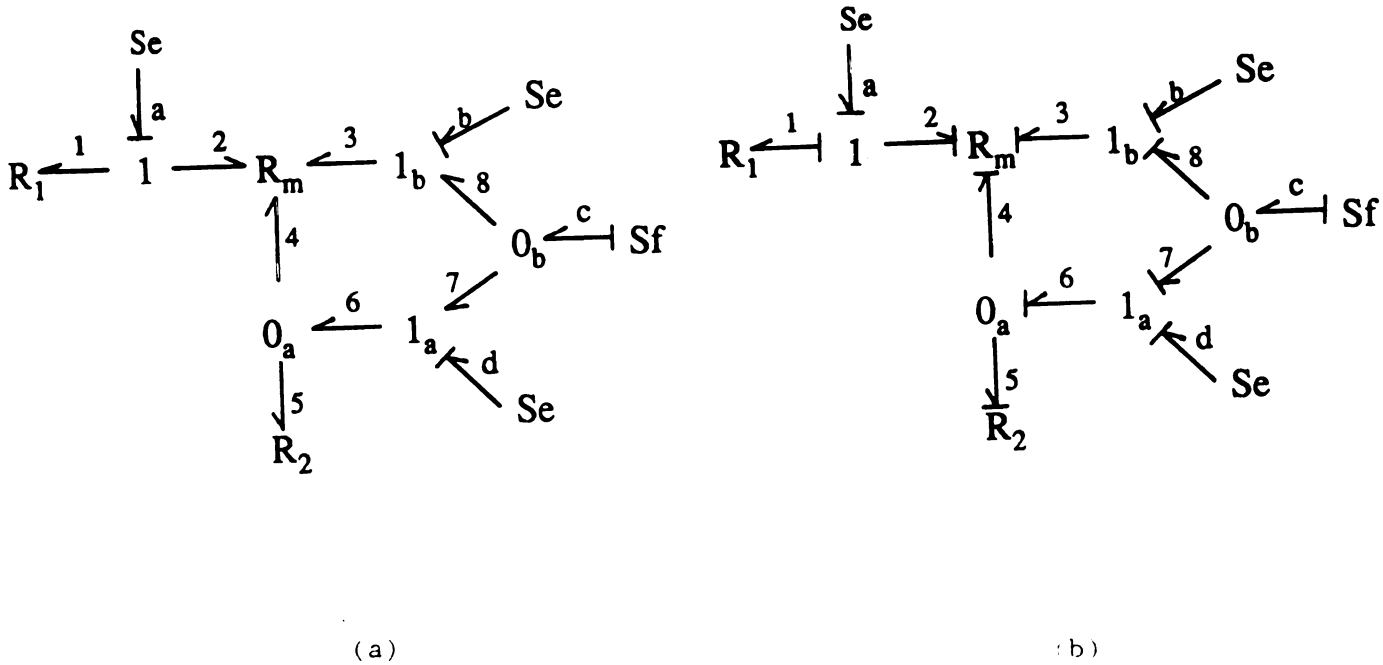


Figure 2-22 Example 1

The number of ports of the multiport R node in this R-block is 2. Since $M_1 = 2 > E_1 = 1$, and $M_1 = F_1$, this R-block belongs to type 2. We have $F_r = M_1 - E_1 = 1$, and $E_g = M_1 - F_1 = 0$. It means that this multiport R node may be in complete G-form, but not in complete R-form. The possible minimum number of the iteration variables is determined to be:

$$N = \min \{2, 3\} + \min \{0, 1\} = 2 \quad (2.60)$$

Based on this result, the multiport R node is assigned to have complete G-form causality (Figure 2-22b). The associated input and output vectors are identified and the constitutive and connective equations are listed as bellow:

$$e_{ro} \quad e_1 = \phi_1(f_1) \quad (2.61a)$$

$$e_5 = \phi_5(f_5) \quad (2.61b)$$

$$e_{rl} \quad \dots$$

$$f_{go} \quad \dots$$

$$f_{gl} \quad f_2 = \phi_2(e_2, e_3, e_4) \quad (2.62a)$$

$$f_3 = \phi_3(e_2, e_3, e_4) \quad (2.62b)$$

$$f_4 = \phi_4(e_2, e_3, e_4) \quad (2.62c)$$

$$f_{ro} \quad f_1 = f_2 \quad (2.63a)$$

$$f_5 = f_c - f_3 - f_4 \quad (2.63b)$$

$$f_{rl} \quad \dots$$

$$e_{go} \quad \dots$$

$$e_{gl} \quad e_2 = e_a - e_1 \quad (2.64a)$$

$$e_3 = e_b + e_d - e_5 \quad (2.64b)$$

$$e_4 = e_5 \quad (2.64c)$$

Since $M_1 = 3$ in this case, we have $M_1 > E_1$ and $M_1 > F_1$. This R-block is of type 2 and $E_r = 3 - 2 = 1$, $F_g = 3 - 2 = 1$. It tells that the multiport R node can not be assigned in either a complete R-form or G-form. The possible minimum number of iteration variables is determined by:

$$N = \min \{3, 3\} + \min \{1, 1\} = 4 \quad (2.67)$$

Assigning causality according to a maximum R-form scheme, the causal graph is shown in Figure 2-23b. Note here that we may also use the maximum G-form scheme in this problem. The constitutive equations are given and the connective equations are derived from the causal graph:

$$D_o: e_{ro} \quad \text{----}$$

$$e_{r1} \quad e_2 = \phi_2(f_2, f_4, f_5, e_3) \quad (2.68a)$$

$$e_4 = \phi_4(f_2, f_4, f_5, e_3) \quad (2.68b)$$

$$e_5 = \phi_5(f_2, f_4, f_5, e_3) \quad (2.68c)$$

$$f_{go} \quad f_1 = \phi(e_1) \quad (2.69a)$$

$$f_6 = \phi(e_6) \quad (2.69b)$$

$$f_{g1} \quad f_3 = \phi_3(f_2, f_4, f_5, e_3) \quad (2.70)$$

$$D_i: f_{ro} \quad \text{----}$$

$$f_{r1} \quad f_2 = f_d - f_1 \quad (2.71a)$$

$$f_4 = f_b - (f_c - f_3) - f_6 \quad (2.71b)$$

$$f_5 = f_c - f_3 \quad (2.71c)$$

$$e_{go} \quad e_1 = e_2 \quad (2.72a)$$

$$e_6 = e_4 \quad (2.72b)$$

$$e_{g1} \quad e_3 = e_e + e_5 - (e_a - e_4) \quad (2.73)$$

The iteration variables are f_3 , f_1 , f_6 , and e_3 and the iteration sequence order is (2.71), (2.68), (2.72), (2.73), (2.70), (2.69). The result can be extended to a IRF containing several multiport R nodes if some modifications are made.

2.5 IRFs with General Junction Structures

2.5.1 General Junction Structures

The junction structure of a bond graph is that portion which represents the power-conserving features of the modeled system; as such it can be viewed as a multiport transformation. In an abstract sense, the junction structure represents the energy topology of a bond graph in the same way that a generalized linear graph represents the topology of an electrical network. Many researchers have made investigations into the properties of junction structures (Karnopp, 1969; Rosenberg, 1971, 1978, 1979; Ort and Martens, 1973; Perelson, 1975). Many system properties, such as solvability and basis order, are obtained by studying the junction structures (Rosenberg and Andry, 1979; Rosenberg, 1980).

Junction structures are classified as simple (SJS), weighted (WJS), or general (GJS), according to whether their junction elements are in the set $\{0 \text{ and } 1\}$, $\{0, 1, \text{ and } TF\}$ or $\{0, 1, TF, \text{ and } GY\}$, respectively. The properties of SJS are contained in WJS, so we shall not discuss SJS further.

The basis order of a WJS is the number of effort variables (E) and flow variables (F) that must be specified for a particular WJS to have all its power variables known. The pair of (E, F) is a unique property of a WJS. The principal application of the basis order is to determine, in advance of assigning causality, the number of independent port effort and flow variables for a given junction structure. The basis order

formula for a given SJS or WJS has been derived by Rosenberg and Moultrie (1980).

The existence of gyrators in a junction structure generally makes the basis order pair no longer unique. The transformation between input and output vectors is not in the form of effort-to-effort and flow-to-flow, since gyrators transform effort to flow (or flow to effort) simultaneously on their ports. The results obtained in the previous sections may not be applied to IRFs containing gyrators. Existing knowledge about the properties of the general junction structure is not complete. In order to extend the results for minimum iteration variables to an IRF with gyrators, we must develop some additional properties of the general junction structure. Until now there has been no general theory that treats the GJS systematically. In this section we develop the basis properties of the GJS.

First we introduce the gyrograph, derived from a GJS in an abstracted form. Then we focus our attention on the relation between causality in bond graph and the graph matching concept in the gyrograph. Following this, we present an algorithm for determining the existence, the F-minimum, and the F-maximum of the basis order for a GJS. Finally, some examples are given.

2.5.2 Gyrographs and Maximum Matching

2.5.2.1 Gyrographs

A gyrobondgraph (CBG) is a bond graph formed from a limited set of multiport nodes, called the primitive set. One such set is {1, GY, Se,

I, R). See Appendix A for details. The junction structure of any bond graph can be represented in GBG terms by the use of 1 junctions and GY nodes. Although the GBG represents a canonical graph form, its essence is the same as that of the standard (9 node type) bond graph. The smaller primitive element set may facilitate the study of some properties of bond graphs. For our purposes, however, it does not seem to be a suitable graph representation for applying the well-developed linear graph theory and associated algorithms to explore the properties of general junction structures. Therefore we consider a more abstract form of the GBG, called a gyrograph (GG) (Paynter, 1967; Zeid, 1982). The GG preserves all the information of the GBG and therefore all the information of the original bond graph. However, it depicts the topological information and connectedness more clearly for certain purposes. Now let us briefly define the gyrograph.

Nodes of a GG are represented by squares and circles. Edges of a GG are lines. The square nodes are used to mark multiport field elements, such as Se, I, and R. These nodes are called environment nodes. The circles represent 1-junctions. An edge represents a gyrobond connector (i.e., 2 bonds and a GY) between a pair of 1-junctions if it joins two circles. An edge represents a bond if it joins a square and a circle. The former is called a junction edge. The latter is called an environment edge. The junction structure is represented by the circles and the edges; its ports are the environment edges. See Appendix A for details.

In a GG the set of environment nodes (squares) is denoted by V_{En} . The set of junction nodes (circles) is denoted by V_J . The set of environment edges is denoted by E_{En} . The set of junction edges is

denoted by E_J . The nodes of the GG are denoted by V , where $V = V_{En} + V_J$. The edges of the GG are denoted by E , where $E = E_{En} + E_J$. No edges are allowed between the environment nodes, since there are no ports defined between elements Se , Sf , C , I , R .

The example in Figure 2-24 illustrates the derivation of a gyrograph from a standard bond graph model. Consider the R-field of Figure 2-24a, which has been isolated from a system. Parts b and c show the associated gyrobondgraph and gyrograph, respectively. The nodes and edges are labeled to show the corresponding elements as the transformations are made from the SBG to the GBG to the GG.

2.5.2.2 Causality and gyrographs

Standard bond graphs and gyrobondgraphs have the same causality properties and can be assigned by the sequential causality assignment procedure. For the primitive set (I , R , Se , I , and GY) we have the following causal properties:

Se has the effort directed outward;

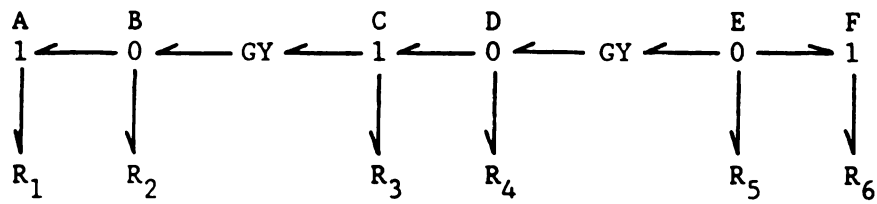
I has two types of causality, integral and derivative;

R has two types of causality, resistance and conductance;

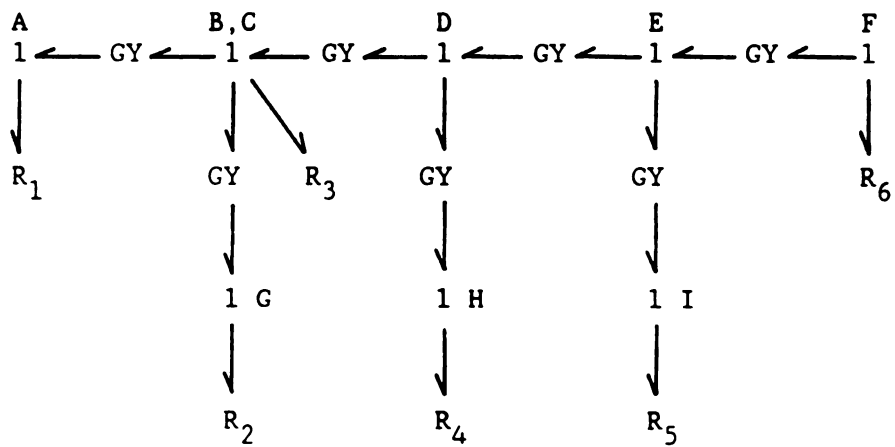
GY has two permissible causal forms, namely, both efforts directed inward or both efforts directed outward; and the

1-junction has exactly one flow directed inward.

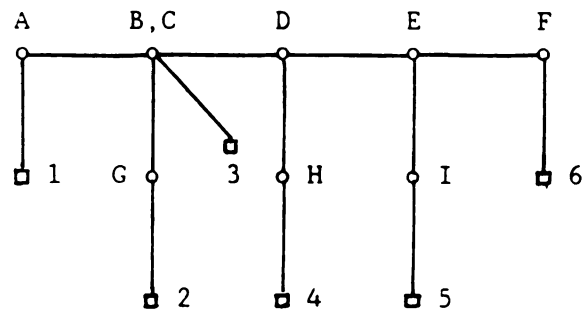
The complete causality assigned to a gyrobondgraph may be recognized in the associated gyrograph by the marking on the edges. For edges in E_J , a marked edge indicates a causal orientation such that both



(a)



(b)



(c)

Figure 2-24 Graphs of an R-field.

(a) The standard bond graph.

(b) The gyrobondgraph.

(c) The gyrograph.

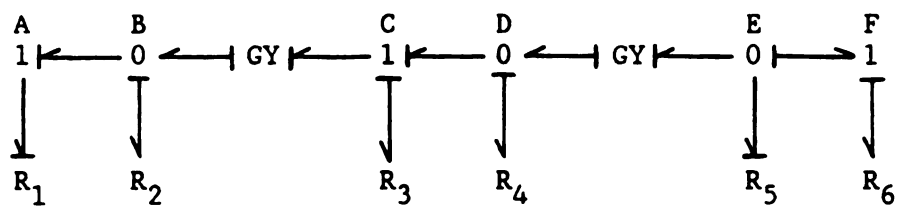
efforts are directed as inputs to the gyrator (or equivalently, both flows are directed as inputs to their adjacent 1-junctions). An unmarked E_j edge of a causal GG indicates the opposite causal state. For edges in E_E , a marked edge indicates that the flow is input to the circle (1-junction), and the effort is input to the square (environment node). An unmarked E_E edge of a causal GG indicates the opposite state.

Figure 2-25 shows causality and its equivalent in three related graphs: the SBG of Figure 2-24, GBG, and its GG. We start with a complete causality for the SBG in part (a). Part (b) shows the corresponding GBG causality, and part (c) shows the equivalent causal information in the causal (i.e., marked) GG. Observe that each circle node in the GG has exactly one marked incident edge.

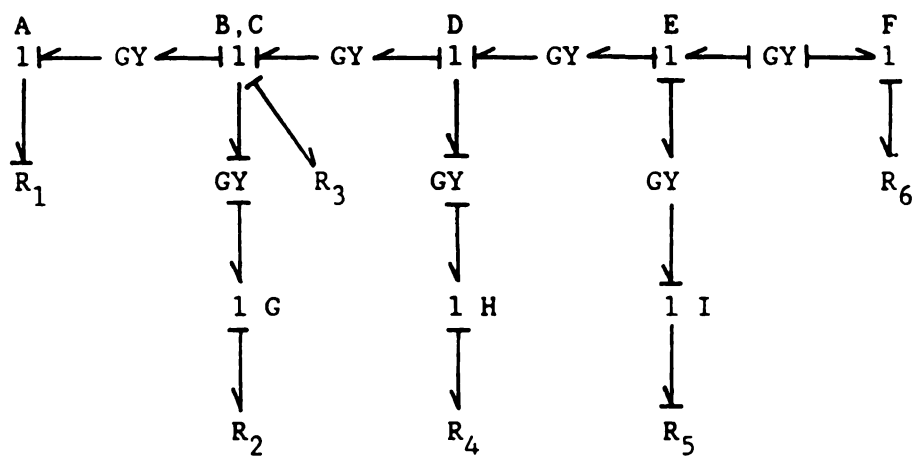
2.5.2.3 Causal GGs and the cardinality matching problem

From the pattern of the marked edges in a gyrograph, we shall refer to an edge-marked GG as a causal GG, since it can be interpreted as a causal GBG. A causal GG also leads itself to interpretation in terms of the cardinality matching problem of standard graph theory. Briefly, the cardinality matching problem (Syslo, et al. 1983) may be stated as: In a given graph, find a maximum matching, that is, a matching with as many edges as possible. The concept of matching is defined next below.

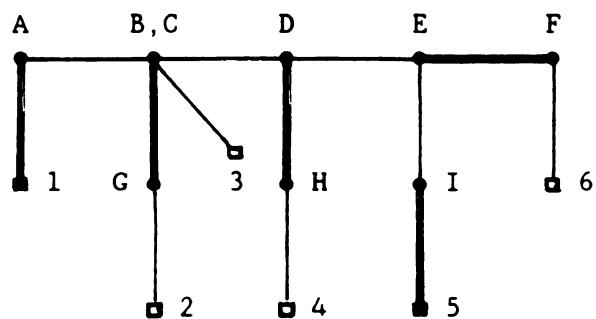
Consider a graph G defined by a set of vertices V and a set of edges E . By definition, a set of edges M in an undirected graph $G = (V, E)$ is called a matching if no two edges in M have a node in common. For example, in Figure 2-26 the edge set $\{(v_1, v_6), (v_2, v_4)\}$, shown in heavy lines, is a matching, because the two edges are not incident to a



(a)



(b)



(c)

Figure 2-25 Causality in different graphs
 (a) in a standard bond graph.
 (b) in a gyrobondgraph.
 (c) in a gyrograph.

common node. The set $\{(v_1, v_6), (v_2, v_5), (v_3, v_4)\}$ is another matching. But $\{(v_2, v_3), (v_3, v_4)\}$ is not a matching because the edges share a node, namely v_3 . A single edge in a graph without self-loops is obviously a matching.

Clearly, a maximum matching in a graph with n nodes can not have more than $(n/2)$ edges. It may have fewer edges. With respect to a given matching M in a graph G , an edge is said to be matched if it is in M . Similarly, a node x is said to be matched or saturated if it is an end node of some matched edge, say, (x, y) . A node that is not matched is called an exposed or free node. In Figure 2-26 nodes v_3 and v_5 are exposed and the rest are saturated. A path $P = (v_1, v_2, \dots, v_k)$ is called an alternating path with respect to a given matching M if the edges of P are alternately in M and not in M . An alternating path that begins at an exposed node and ends at another exposed node is called an augmenting path. The maximum matching algorithm is based on Berge's Theorem that a matching M in a graph G is maximum if G has no augmenting path with respect to M (Syslo, et al., 1983).

For our later purpose, we introduce the concept of an alternating cycle. An alternating cycle is an alternating path that begins at an exposed node and ends at the same node. For example, in Figure 2-26 the path (v_3, v_2, v_4, v_3) is an alternating cycle with respect to matching $M = \{(v_1, v_6), (v_2, v_4)\}$. So are the paths (v_5, v_4, v_2, v_5) and $(v_5, v_4, v_2, v_1, v_6, v_5)$.

For a 1-junction node, exactly one incident bond must be set with a flow input to the 1-junction. The implication of this requirement in a gyrograph is that a circle node must have exactly one marked incident

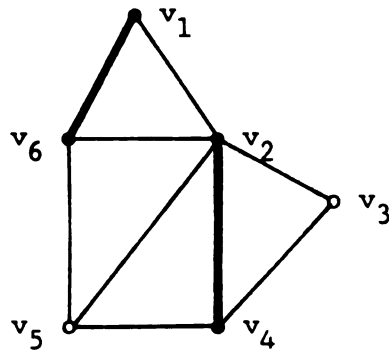


Figure 2-26 A graph and one of its matchings

edge. In a gyrograph marked with the equivalence of complete causality, all the circle nodes must have exactly one marked incident edge. Hence they are all saturated. This property will be utilized to explore the basis order of a GJS.

2.5.3 Basis Order Algorithm for General Junction Structures

From Equations 2.23a and b we see that the basis variable structure (E and F pair) for a given WJS is unique. However, for a junction structure with essential gyrators (Rosenberg, 1979; Breedveld, 1984) the basis variable structure is not unique. That is, the number of effort variables (E) and the number of flow variables (F) required as inputs to determine the general junction structure is not a unique pair.

We here present an algorithm for determining the existence of basis order properties for an arbitrary GJS. If a basis does exist, then we

can determine the minimum and the maximum flow basis orders, $F_{\min}$ and $F_{\max}$, respectively. We will employ the gyrograph and apply a variation of the standard cardinality matching algorithm to find the properties of the basis order.

In the study of the junction structure, we will sometimes focus our attention to the portions of the gyrograph which contains only junction nodes and edges between them; namely, subgraphs induced by junction node sets. The following definitions are introduced to assist in development.

Definition 1. External junction node: A junction node which is adjacent to at least one environment node is called an external junction node. The set of external junction nodes is denoted as V_E .

Definition 2. Internal junction node: A junction node which is adjacent to no environment nodes is called an internal junction node. The set of internal junction nodes is denoted as V_I .

Definition 3. Junction gyrograph: A subgraph of a gyrograph in which all environment nodes and their incident edges are removed is called a junction gyrograph. We use the symbol G_J to denote a junction gyrograph.

For a given gyrograph to have a basis each junction node must exactly have one marked incident edge. Such nodes are said to be saturated. This corresponds to each of the 1-junctions in the associated GBG having an incident bond which gives the junction a strong causal determination. The basis order properties are derived from the junction gyrograph G_J and the gyrograph GG .

The nodes in a junction gyrograph G_J may be classified as belonging to either the external set V_E , or the internal set V_I . In a marked gyrograph the nodes may be saturated or free. The collection of free nodes in the internal set and external set are denoted as V_I^f and V_E^f , respectively. The collection of saturated nodes in the internal set and the external set are denoted as V_I^s and V_E^s , respectively.

Let us study the junction gyrograph first. Suppose we have a maximum matching for a given junction GG. There are four cases (not all mutually exclusive) that can exist. We present each case and discuss its implications.

Case 1. All the nodes of internal set V_I are saturated. Then a basis solution exists.

Case 2. There exists an alternating path P which starts at a free node in the internal set V_I and ends at a saturated node in the external set V_E . Then we reverse the marking of the edges (i.e., the edges of P currently in matching M are removed from the matching and those edges of P not in M are put into the matching). The size of the matching M is unchanged and so is the number of saturated nodes. However, the starting node in V_I is now saturated, while the ending node in V_E becomes free. By repeating this procedure we can move all such free nodes from V_I^f to V_E^f . An example of this case is shown in Figure 2-27a. It depicts a portion of the maximum matching generated by a maximum matching algorithm. The matched edges are shown in heavy lines. Along the alternating path (v, x, y, z, u) , node v is a free internal node, whereas u is a saturated external node. This path can be changed to that of Figure 2-27b by reversing the markings of the edges. Thus the node v

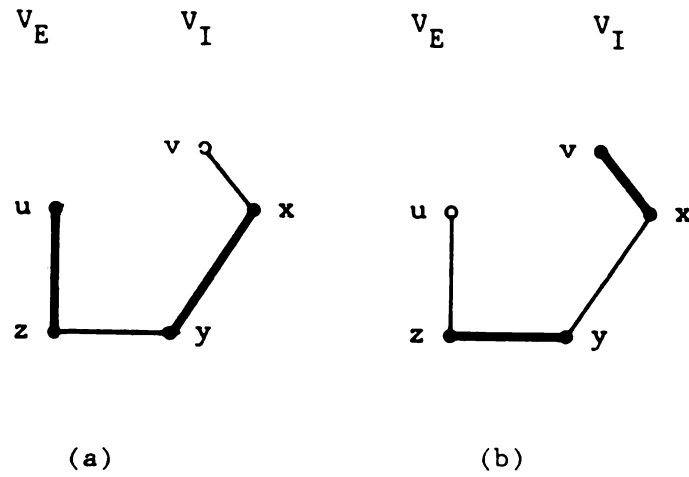


Figure 2-27 An alternating path.
 (a) Before reversing markings.
 (b) After reversing markings.

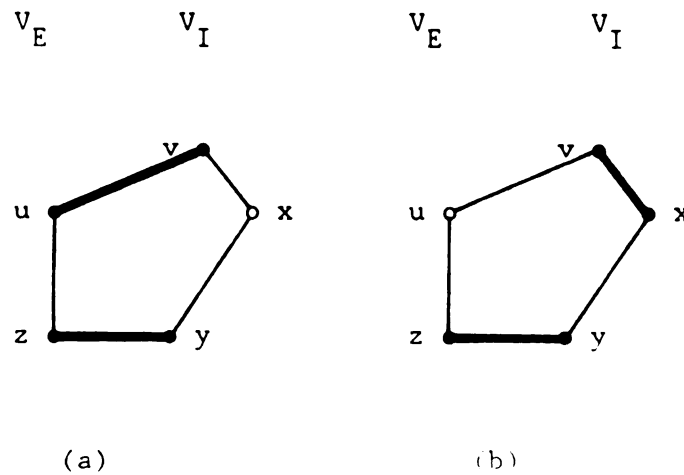


Figure 2-28 An alternating cycle.
 (a) Before reversing markings.
 (b) After reversing markings.

becomes a saturated node, while the node u becomes a free node.

Case 3. There exists an alternating cycle C , which starts and ends at a free node v in the internal set V_I^f and contains at least one external node u in V_E . Then we can reverse the markings of some of the edges of C to make v saturated and u free. This procedure does not change the size of the matching set M or the number of free nodes, but it moves the position of the free node from the internal set V_I to the external set V_E . This is illustrated in an example (Figure 2-28). Suppose that after a maximum matching is generated, an alternating cycle $C(x, y, z, u, v, x)$ is identified. There is a free node x in V_I . The cycle contains nodes z and u which belong to the external set V_E . If we choose node u to be the destination for the free node x , we reverse the markings starting at the marked edge (u, v) along the partial cycle (u, v, x) until edge (v, x) , such that node x becomes saturated and node u becomes free (Figure 2-28b). We may also choose node z to be the destination of the free node x , by reversing another partial cycle (z, y, x) starting at the marked edge (z, y) . From this example, we can extract the fact that any saturated node of V_E in the alternating cycle can be the destination of an internal free node in the same cycle.

Case 4. There exist one or more free nodes which may not be saturated by means of alternating paths and/or cycles. This means that the 1-junction associated with the free node can not be assigned by deterministic causality. Hence no solution exists for this GJS.

The full saturation of the set V_I implies the completeness of the deterministic causality assignment to the internal junction nodes, since every node has one flow input. The free nodes in the set V_E may get

saturated by their adjacent environment nodes. Therefore the existence of a solution to the GJS can be judged by whether $|V_I^f| = 0$ or not. ($| \cdot |$ denotes the size of a set.) If an external node is saturated, then it implies that the associated 1-junction has been assigned a flow input. Thus the adjacent port bond must have effort as input to the 1-junction. If an external junction node is free, then the adjacent port bond has flow as input to the 1-junction. Since a maximum matching algorithm is applied, the number of free nodes in V_E is a minimum. Hence a maximum matching with all free nodes in set V_E gives the minimum number of flow inputs to the associated GJS, since each external node has at least one environment edge (i.e., each external 1-junction has at least one port).

Now let us expand the junction gyrograph to include all the environment nodes. Starting with the maximum matching obtained within the junction gyrograph, a second maximum matching with respect to the entire gyrograph may be done such that all the nodes in the set V_E are saturated. In this process only two cases occur. First, a free node in the set V_E resulting from the first maximum matching can be saturated by simply marking the incident environment edge. Thus the adjacent environment node becomes saturated. Since one free external junction node makes one environment node saturated, $F_{\min}$ free external junction node will make $F_{\min}$ saturated environment nodes. In the second case, an augmenting path starting at an environment node and ending at another environment node is found. Reversing matching of this path will increase the size of the matching set M by one and the size of saturated environment nodes by two. If a maximum matching is found, then the number of saturated environment nodes will be a maximum. It corresponds the maximum number of flow input variables; namely, the F-maximum basis.

An algorithm for determining the existence, the F-minimum basis, and the F-maximum basis for general junction structures is stated based on the above discussion.

Basis Order Algorithm for GJS

Assume an unmarked GG derived from a given GJS.

(1) Identify the sets V_E , V_I , and V_{En} .

(2) Start with the junction gyrograph G_J induced by the node set V_E and V_I .

(3) Apply the Pape-Conrad maximum matching algorithm (Syslo, et. al., 1983) to the junction gyrograph G_J .

* If ($|V_I^f| = 0$) then
a solution exists.

Else

Start at a free node $v \in V_I^f$.

If (there exists an alternating path which ends at a node
 $u \in V_E$) then

reverse the edges. Therefore we have

$$\begin{aligned} V_I^f &= V_I^f - v, & |V_I^f| &= |V_I^f| - 1; \text{ and} \\ V_E^f &= V_E^f + u, & |V_E^f| &= |V_E^f| + 1. \end{aligned}$$

Go to (*) above.

Elseif (there exists an alternating cycle which contains a
node $u \in V_E$) then

reverse the markings of edges starting from the matched
edge incident with u to the edge incident with v . Therefore
we have

$$V_I^f = V_I^f - v, \quad |V_I^f| = |V_I^f| - 1; \text{ and}$$

$$V_E^f = V_E^f + u, \quad |V_E^f| = |V_E^f| + 1.$$

Go to (*) above.

Else

No solution exists. Stop.

Endif

Endif

(4) $F_{\min} = |V_E^f|$. The minimum number of flow inputs to this general junction structure is equal to the number of free nodes in the external set.

(5) Apply the Pape-Conradt maximum matching algorithm to the entire gyrograph GG.

(7) $F_{\max} = |V_{En}^s|$. The maximum number of flow inputs to this general junction structure is equal to the number of saturated environment nodes.

All bases for a given GJS have F ranging between $F_{\min}$ and $F_{\max}$. From the discussion on the maximum matching for the entire gyrograph we also deduce that

$$F = F_{\min} + 2k \quad (2.74)$$

where $k = 0, 1, 2, \dots, K$ and K is given by $K = (F_{\max} - F_{\min}) / 2$. If the $F_{\min}$ and the $F_{\max}$ of a general junction structure obtained by the above algorithm are the same, it indicates that this GJS has unique basis order numbers (E, F) .

2.5.4 Examples

2.5.4.1 Example 1

A bond graph with the environment nodes denoted by EN is given in Figure 2-29a. Its associated gyrobondgraph is identical, and the gyrograph is given in Figure 2-29b. The nodes a, b, c and d, in the gyrograph correspond to the junctions 1a, 1b, 1c and 1d in the bond graph (gyrobondgraph), respectively. The junction gyrograph G_J consists of nodes a, b, c, and d. Nodes a, b, and c belong to the external set and node d belongs to the internal set. Applying the maximum matching algorithm to this gyrograph yielded no free node in the external set. Thus $F_{\min} = |V_E^f| = 0$. By using the result from the maximum matching in G_J , we applied the maximum matching algorithm again to the entire gyrograph (Figure 2-29d). There was one augmenting path $\{(A, a), (a, c), (c, C)\}$. Reversing this path yielded a augmented matching set $M = \{(A, a), (b, d), (c, C)\}$ (Figure 2-29e). Two environment nodes are saturated. We found that $F_{\max} = 2$. The corresponding causality assignments for $F_{\min}$ and $F_{\max}$ are shown in part f and g. Notice that the difference between the $F_{\min}$ and the $F_{\max}$ in this example is 2.

2.5.4.2 Example 2

An arbitrary gyrograph is given in Figure 2-30a. In this graph, $V_E = \{a, b, c, d\}$, $V_I = \{e, f, g, h, i, j, k\}$, and $V_{En} = \{A, B, C, D, E, F, G\}$. We first apply the maximum matching algorithm to its junction gyrograph G_J which is induced by the set V_E and V_I . A maximum matching may be found, namely, $M = \{(b, f), (c, g), (d, h), (e, i), (j, k)\}$. Since $V_I^f = \{0\}$, a solution does exist (Figure 2-30b). The F -

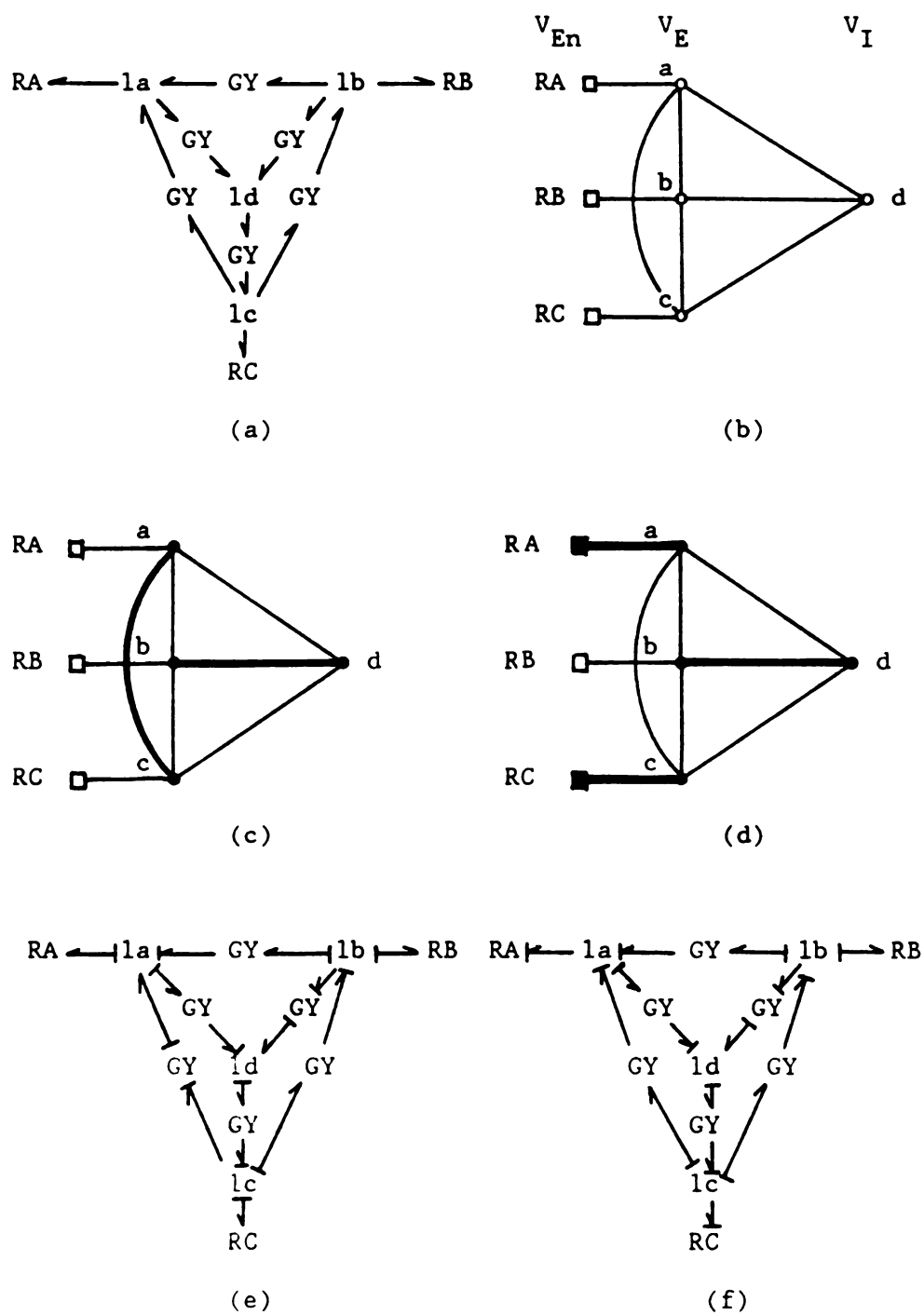
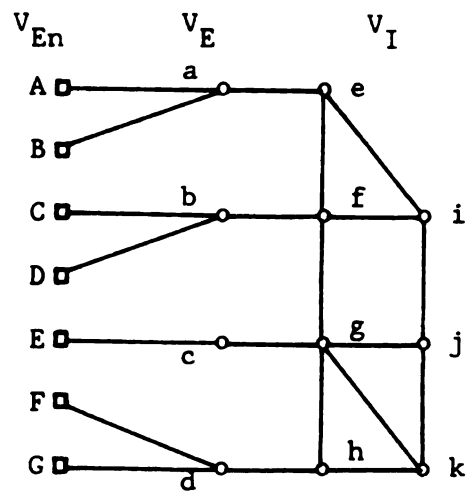
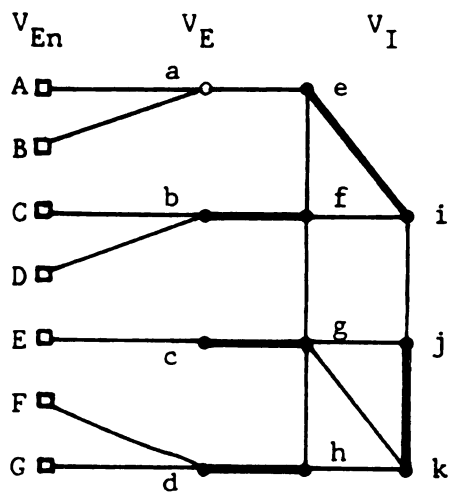


Figure 2-29 Example 1

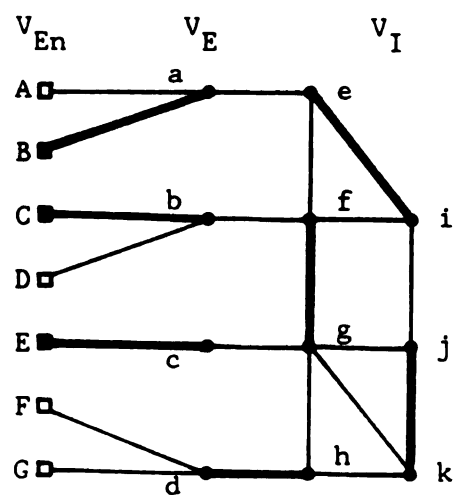
- (a) Bond graph and gyrobondgraph
- (b) Gyrograph
- (c) Maximum matching in G_j
- (e) Maximum matching in GG
- (e) Causality corresponding to the F-minimum basis
- (f) Causality corresponding to the F-maximum basis



(a)



(b)



(c)

Figure 2-30 Example 2.

(a) Gyrograph

(b) Maximum Matching of G_J (c) Maximum Matching of GG

minimum basis was found from $F_{\min} = |V_E^f| = 1$. Now we apply the maximum matching algorithm to the entire gyrograph. An augmenting path is found; namely, $P = ((C, b), (b, f), (f, g), (g, c), (c, E))$. Reversing the markings of the path made the environment nodes C and E saturated. The number of saturated environment nodes were increased by two. Thus the F-maximum is found to be $F_{\max} = |V_E^s| = 3$. In this case the flow inputs will be applied to the junctions a, b, and c. Note that we may also find another solution (a, b, d). Also we noticed that $F_{\max}$ is greater than $F_{\min}$ by 2 in this case.

2.5.5 Remarks on the Algorithm

An algorithm for determining the basis order properties of an arbitrary GJS has been developed through the use of the gyrograph representation and the graph matching concept. The existence of a basis order is determined by the absence of free nodes in the internal set of the junction gyrograph. The number of free nodes in the external set of the junction gyrograph gives the F-minimum basis order of the GJS. The resulting marked gyrograph is further used to determine the F-maximum basis order by increasing the number of saturated nodes in the environment set. The possible number of flow inputs to a GJS ranges between $F_{\min}$ and $F_{\max}$ in steps of 2. We conjecture that if $F_{\min} = F_{\max}$ in a GJS, no essential gyrators exist in the GJS. This remains to be investigated.

The methodology for determining the basis order of a bond graph with a general junction structure by using a transformed gyrograph may be applied to the solution of the implicit R-field problem and the implicit C(I)-field problem. The identification of the basis properties

of the input variables and determination of feasible inputs can contribute to improving computational efficiency for large scale nonlinear dynamic systems.

MODEL ORDER REDUCTION METHODS

A simplified model of engineering system is always highly desired in dynamic analysis, synthesis, and design. The simplification of a dynamic system model brings many benefits to engineers. First, repetitive simulations become easier and cheaper to perform, since the computational load due to large dimension size and the widely separated system time constants may be reduced dramatically. It happens in the investigation of the influence on the system performance as some of system parameters have been varied. Second, the complexity of a higher order model often makes it difficult to obtain a good understanding of the behavior of the system. Salient features of the system, previously hidden in a mass of detail, may be revealed. Third, controllers may be designed for the reduced model, since most currently available control design methods only work on small-dimension systems.

The simplification of a system model can be achieved by reducing the model order, or by eliminating the adverse parameters which cause the computational difficulties. Because of their importance in system analysis and the design of controllers, model order reduction methods have received considerable attention over the past three decades. In the existing literature there are a number of books dedicated to this topic (Decarlo and Saeks, 1981; Happ, 1971; Jamishidi, 1983; Michel and Miller, 1977; Sage, 1978; Siljak, 1978; Kokotovic, Khalil, and O'Reilly, 1986) and numerous research papers. The objective of model order reduction is to find a lower order model which preserves the dynamics of more complex high order system in both time and frequency domains. The

techniques in the existing literature can be divided into two groups. The first group of methods attempts to retain the dominant modes of the original systems. It includes all aggregation methods and perturbation methods. Another approach is based on applying an identification procedure to input-output data obtained by driving the original system with a specific input, for example, Walsh functions (Kawaji and Shiotsuki, 1985). Since the latter is not widely applied in practice, we will restrict ourselves to the first group. All the methods surveyed below are concerned with time-domain models. Alternatively, the linear time-invariant systems in state-space form can be represented in frequency domain. By far the greatest effort in model order reduction techniques based on frequency domain has been for single-input single-output systems. We will not discuss these techniques here and in this regard Jamishidi's book (1983) would be an excellent reference.

3.1 Aggregation Methods

The notion of aggregation was introduced in the control literature by Aoki (1968). The intuitive idea behind the notion of aggregation is quite simple. Suppose that S_1 is a mathematical description of a physical system using a given set of variables, and S_2 is a consistent description of the same system using smaller set of variables. Then S_2 is termed an aggregate model for S_1 and the variables of the system S_2 are termed aggregate variables.

In the literature there are a number of aggregation methods. The most basic aggregation method is the exact aggregation which illustrates the notion of aggregation most clearly.

Consider the system

$$\dot{x}(t) = Ax(t) + Bu(t) \quad (3.1a)$$

$$y(t) = Dx(t) \quad (3.1b)$$

Let $z(t) = Cx(t)$, where x is a vector of dimension of n and z is a vector of dimension of r . We want to obtain a new system model with lower dimension r , i.e.

$$\begin{aligned} \dot{z}(t) &= Fz(t) + Gu(t) \\ &= FCx(t) + Gu(t) \end{aligned} \quad (3.2a)$$

$$y(t) = Wz(t) \quad (3.2b)$$

By the requirement of consistency(or dynamic exactness), for any $u(t)$ and $z(0) = Cx(0)$, we need

$$CA = FC \quad (3.3a)$$

$$CB = G \quad (3.3b)$$

$$\text{and } WC \approx D \quad (3.3c)$$

If the above equations are satisfied, then the triple (F,G,W) is said to be a perfect aggregation of the triple (A,B,D) relative to the aggregation matrix C .

Further insight into the nature of the class of matrices for which dynamic exactness can be achieved is obtained by realizing that the aggregation problem as posed for linear system is in fact a problem of minimum realization. The exactness of the original system and the

reduced system only can be achieved when there are pole-zero cancellation in the transfer function of the original system.

However, the aggregate state variables $z_i(t)$ will not in general correspond exactly to physical variables (Sandell, et al., 1978). Therefore, an alternative point of view, which may be more useful in applications, is to regard $z(t)$ as an approximation to physical variables. In other words in addition to (3.3a) and (3.3b), we desire

$$z(t) \approx y(t) = Cx(t) \quad (3.4)$$

where the matrix C picks out components or linear combinations of components of $x(t)$ that are to be approximated. Of course, the choice of the aggregate matrix C can greatly influence the nature of the approximation.

There are many other aggregation methods such as controllability matrix approach (Aoki, 1968), continued fraction method (Chen and Shieh, 1969), chained aggregation (Tse, et al., 1977), aggregation via covariance equivalent realization (Yousuff, et al., 1985) and model reduction via balanced state-space representation (Pernebo and Silverman, 1982). All these methods are based on matrix similarity transformation, therefore, they are not directly applicable to the general nonlinear problems.

3.2 Modal Method

The modal method (Davison, 1967) is a well-known model order reduction method. Consider a generalized linear system in the state-space form

$$\dot{x}(t) = Ax(t) + Bu(t) \quad (3.5a)$$

$$y(t) = Cx(t) \quad (3.5b)$$

Assume that A has distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ with negative real part and M is the modal matrix with columns consisting of the corresponding eigenvectors of A . Define a new state vector $\eta(t)$ which is transformed via

$$\eta(t) = M^{-1}x(t) \quad (3.6)$$

then the system equation can be transformed into

$$\eta(t) = \begin{vmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ & & \ddots \\ 0 & & \lambda_n \end{vmatrix} \eta(t) + \begin{vmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{vmatrix} u(t) \quad (3.7a)$$

$$y(t) = [c_1, c_2, \dots, c_r]^T \eta(t) \quad (3.7b)$$

Suppose that the eigenvalues cluster into dominant and undominant eigenvalues and that the undominant eigenvalues are asymptotically stable. If there exist some b_i 's = 0, we can say that the corresponding states are not controllable, and if there exist some c_j 's = 0, we can

say that these corresponding states are not observable. From input-output consideration, the only important modes are those which are both controllable and observable. Therefore, we may drop the equations corresponding to the i 's and j 's. Eliminating the uncontrollable and unobservable modes leads to a minimal realization. By the same reasoning, we may further approximate the transformed system by dropping the states which are weakly controllable or weakly observable, that is, the corresponding b_i 's and c_j 's are small.

Since the transformed system matrix has diagonal form, the new system states are decoupled. We partition η according to the the dynamic speed, controllability and observability into subblocks as the perturbation form:

$$\dot{\eta}_1(t) = J_1 \eta_1(t) + G_1 u(t) \quad (3.8a)$$

$$\epsilon_2 \dot{\eta}_2(t) = J_2 \eta_2(t) + G_2 u(t) \quad (3.8b)$$

$$\dot{\eta}_3(t) = J_3 \eta_3(t) + \epsilon_3 G_3 u(t) \quad (3.8c)$$

$$\dot{\eta}_4(t) = J_4 \eta_4(t) + G_4 u(t) \quad (3.8d)$$

$$y(t) = H_1 \eta_1(t) + H_2 \eta_2(t) + H_3 \eta_3(t) + \epsilon_4 H_4 \eta_4(t) \quad (3.8e)$$

If we set $\epsilon_i = 0$, $i = 2, 3, 4$, then

$$\dot{\eta}_1(t) = J_1 \eta_1(t) + G_1 u(t) \quad (3.9a)$$

$$y(t) = H_1 \eta_1(t) - H_2 J_2^{-1} G_2 u(t) \quad (3.9b)$$

This method is workable theoretically, but not advised in practice. The computational effort is not trivial for large scale system since the work involved in eigenanalysis of an $n \times n$ matrix goes up as n^3 . The other trouble is from the numerical characteristics of large scale systems which are ill-conditioned mathematically. Besides, if the

state variable of the original model $x(t)$ represents physical variables, the physical nature of the variables will be lost since $\eta_1(t)$ represents mathematical variables which in most case do not correspond to any physical variables.

3.3 Lyapunov Function Method

Stability is one of the most important properties of a dynamic system. There are a number of qualitative analysis methods based on this property for large-scale systems. Lyapunov's second method is an ideal mechanism for accomplishing the aggregation plan in the stability analysis of large-scale dynamic systems (Siljak, 1978). Actually, the Lyapunov method itself can be viewed as an aggregation process. A stability property, involving several state variables, is entirely represented by a single variable --- the Lyapunov function. However, this approach simplifies the stability problem, but sacrifices detailed information about the size of variations of each separate state variable.

The concepts of vector differential inequalities and vector Lyapunov function have been developed by Matrosov (1962) and Bellman (1962) and other researchers. The concept associates with a dynamic system several functions (say s) in such a way that each of them determine the desired stability properties in the system space (of dimension $n > s$) wherever the others fail to do so. These scalar functions are considered as components of a vector Lyapunov function, and a differential inequality is formed in terms of this function, using the original system of equations. As in the case of scalar Lyapunov function, the stability properties of an n -th order system are

determined by considering only the s-vector differential inequality Lyapunov functions. This can bring about a considerable reduction in the dimensionality of a stability problem. It should be mentioned immediately that there is no general systematic procedure for choosing vector Lyapunov functions and that is the most serious drawback of the approach.

3.4 Perturbation Methods

The other scheme of model order reduction for large-scale system is perturbation, which is based on ignoring certain interactions of the dynamic or structural nature in a system.

Perturbation methods are useful for dealing with a system that can be approximated by a system of simpler structure. Mathematically, the difference in the response between the actual and approximated systems is modeled as a perturbation term driving the latter. In principle they can be applicable to both linear and nonlinear problems. The perturbations are divided into two classes of regular and singular perturbations.

3.4.1 Regular Perturbation Method

The regular perturbations (Kokotovic, et al., 1969) are those that appear in the right-hand side of a differential equation, the general formulation is the following form

$$\dot{x}_1(t) = f_{11}(x_1) + \epsilon f_{12}(x_2) + b_1 u_1(t) \quad (3.10a)$$

$$\dot{x}_2(t) = \epsilon f_{21}(x_1) + f_{22}(x_2) + b_2 u_2(t) \quad (3.10b)$$

where ϵ is a small positive parameter. The system is connected by the small (weak) connections $\epsilon f_{12}(x_2)$ and $\epsilon f_{21}(x_1)$ and it can be decomposed into two completed independent subsystems by ignoring the weak connections. The computation of the two independent lower dimensional problems is fewer than that of the single high dimensional problem. This effect is enhanced for more than two subsystems.

3.4.2 Singular Perturbation Method

By singular perturbation is meant a perturbation to the left-hand side of a differential equation (Kokotovic, et al., 1986). Consider a dynamic system of the form

$$\dot{x}(t) = f(x, y, t, \epsilon) \quad (\text{slow subsystem}) \quad (3.11a)$$

$$\epsilon \dot{y}(t) = g(x, y, t, \epsilon) \quad (\text{fast subsystem}) \quad (3.11b)$$

where ϵ is a small positive parameter. If we set $\epsilon = 0$, then the reduced slow subsystem becomes

$$\dot{\bar{x}}(t) = f(\bar{x}, y, t, 0) \quad (3.12a)$$

$$0 = g(\bar{x}, y, t, 0) \quad (3.12b)$$

If Equation 3.12b has isolated roots $y = h(\bar{x}, t)$, the limit model for the slow system is

$$d\bar{x}(t)/dt = f(\bar{x}, h(\bar{x}, t), t, 0) \quad (3.13)$$

The fast or boundary layer system can be obtained by stretching the time scale from t to $\tau = (t-t_0)/\epsilon$ as

$$\frac{d\hat{y}}{d\tau} = g(\bar{x}, \hat{y}, t_0 + \epsilon\tau, \epsilon) \quad (3.14)$$

Under the condition that $\text{Re} \lambda(\frac{\partial g}{\partial y}) \leq -\mu < 0$, the Tichonov Theorem can be applied such that

$$x(t) = \bar{x}(t) + O(\epsilon) \quad (3.15a)$$

$$y(t) = \hat{y}(t) + h(\bar{x}, t) + O(\epsilon) \quad (3.15b)$$

Since the concept involves essentially an asymptotic approximation, quantitative design results are difficult to obtain --- it is hard to say 'how small is small enough'. Furthermore, singular perturbation is not a coordinate free concept, which may be the reason for a lack of modeling procedures and prescriptive (computer-oriented) decomposition techniques for model reduction via the characterization of the fast and slow scales (Siljak, 1983).

As Sandell and Athans (1978) state: "From the practical point of view, the main problem with this method is that a model of a physical system is hardly ever given in the standard form with the slow and fast variables separated and the parameter ϵ conveniently appearing in the left-hand side of the equations. It is a completely nontrivial exercise, requiring considerable physical insight, to model a physical system with slow and fast modes in the framework demanded by the theory!" This problem is naturally more severe for a poorly understood large scale system.

3.5 Component Connection Methods

3.5.1 Component Connection Method

For reasons of efficiency it is often profitable to handle the component equations of a system as separate entities. This permits one to store the different component models separately in computer memory and to analyze them one at a time. By decomposing a large scale system into a number of smaller subsystems (components) and connecting them based on their interactions, the component connection model can be used (Decarlo and Saeks, 1981).

An interconnected dynamical system may be composed of many components, each of which has a mathematical model of the following form

$$dx_i/dt = f_i(x_i, a_i) \quad (3.16a)$$

$$b_i = g_i(x_i, a_i) \quad (3.16b)$$

where a_i is the vector of input signals for the i -th component; b_i is the vector of output signals for the i -th component; and x_i is the state vector of the i -th component.

The interaction between the i -th component and the rest of components is described by an algebraic equation

$$a_i = L_{11i} b + L_{21i} u \quad (3.17)$$

where u is the system input vector. From a theoretical point of

view, however, it is convenient to lump the n component equations together, forming a single Composite Component Model. It takes the form

$$dx/dt = f(x, a) \quad (3.18a)$$

$$b = g(x, a) \quad (3.18b)$$

and

$$a = L_{11} b + L_{12} u \quad (3.19a)$$

$$y = L_{21} b + L_{22} u \quad (3.19b)$$

where x is the composite component state vector, a and b are the composite component input and output vectors, respectively, and u and y are the composite system input and output vectors.

The component connection model divides the system into two sets of equations: component equations, characterized by partially decoupled differential equations, and the connection equation, characterized by coupled linear algebraic equations.

There are two algorithms based on the component connection method. One is the Sparse Tableau Approach which "stacks" the various component equations together with the connection equations to form a large, highly sparse set of simultaneous equations. Given an input vector, u , and a set of initial conditions, one can solve for a , b , and y by use of sparse matrix inversion. The other one is the Relaxation Algorithm, which builds around a predictor-corrector integration scheme. It can solve linear and nonlinear systems. The algorithm allows one to use a different variable order and/or step-size integration routine for each component of the system.

3.5.2 Diakoptics

It should be mentioned that research on decomposition-aggregation methods was conducted by Kron in the 1950s. He developed a scheme called Diakoptics (Kron, 1963) whose main procedure can be summarized as follows

- (a) Tear the system apart into logical groups, each of which can conveniently be analyzed as one unit;
- (b) Set up the equations of each component unit separately, as if the other units were non-existent;
- (c) Set up a "connection matrix" C showing how the various components are interconnected;
- (d) Using the matrix C with the laws of transformation of tensor analysis, establish the equations of the interconnected system;
- (e) Solve the equations piecewise.

It is rare to find a new research paper on Diakoptics in the recent literature. Harrison pointed that "Diakoptics which has been successful in solving large electric networks, turned out to be not as successful in other type of models" (Harrison, 1972).

Karnopp (1970) also stated "All too often, however, the mathematical style of Kron's presentation obscured even his basic philosophy and many workers in the field of system dynamics elected to ignore his contribution. If we examine the relationship of the Diakoptics with bond graphs, a junction structure may be regarded as an almost physical representation of the tensors Kron discussed". However,

Kron's basic philosophy of decomposition-aggregation is still significant to the research on the large scale systems.

3.6 Component Cost Analysis of Large Scale Systems

The performance of a dynamic system is often evaluated in terms of a performance metric V . The performance metric V might represent the system energy or a norm of the output error over some interval of time. A question is "what fraction of the overall system performance metric V is due to each component of the system?" Based on the notion of significance of system component in a dynamic system, the component cost analysis for linear systems has been developed by Skelton, et al., (1980, 1983).

Component cost analysis (CCA) consists of the decomposition of V into the sum of contributions V_i associated with each component state X_i , where the V_i 's satisfy the cost-decomposition property,

$$V = \sum_{i=1}^n V_i \quad (3.20)$$

It seems equally natural and basic, therefore, to characterize the system's behavior in terms of contributions from each of the entities in the system. The CCA algorithm is summarized as:

Step 1. Determine a performance metric

$$V = \lim_{t \rightarrow \infty} E || y ||^2 \quad || y ||^2 = y^t y \quad (3.21)$$

for the system

$$\dot{x}_i = \sum_{j=1}^n A_{ij} x_j + D_i W \quad y = \sum_{j=1}^n C_j x_j \quad (3.22)$$

Step 2. Compute V_i from $V_i = \text{tr}[XC^t C]_{ii}$ and

$$0 = XA^t + AX + DD^t \quad (\text{Ricatti-type equation}) \quad (3.23)$$

Step 3. Rank the component costs in the manner

$$|V_1| \geq |V_2| \geq \dots \geq |V_n|$$

The 'most critical' component of the system is x_1 having component cost V_1 and 'the least critical' component of the system is x_n having component cost V_n .

Step 4. The least critical components will be deleted from the system.

The accuracy is controlled by the cost perturbation index defined by

$$\Delta = \sum_{i=1}^k V_i / \sum_{i=1}^n V_i \quad k \leq n \quad (3.24)$$

where k is the number of retained components.

3.7. Model order reduction in bond graph models

As has been suggested in the brief discussion about Diakoptics above, there is in a bond graph model the potential for exploiting efficient solution techniques. Here we mention two methods, one of which is based on the Sequential Causality Assignment Procedure (Rosenberg and Karnopp, 1983), and is the standard approach for implementing

computational methods based on bond graphs (Van Dixhoorn, 1977; Granda, 1985; Rosenberg, 1984, 1986). The second method is newer and has not been implemented computationally at this time.

3.7.1. Sequential causality assignment method

There is a close connection between the Sequential Causality Assignment Method and the Component Connection Method. When the sequential causality assignment procedure has been followed for a bond graph model, the system equations can be expressed in the form

$$Z_i = f_i(X_i, X_d) \quad (3.25a)$$

$$Z_d = f_d(X_i, X_d) \quad (3.25b)$$

$$D_o = g(D_i) \quad (3.25c)$$

$$U = h(t) \quad (3.25d)$$

$$dx_i/dt = S_{11}Z_i + S_{12}(X_d/dt) + S_{13}D_o + S_{14}U \quad (3.25e)$$

$$Z_d = S_{21}Z_i + S_{24}U \quad (3.25f)$$

$$D_i = S_{31}Z_i + S_{33}D_o + S_{34}U \quad (3.25g)$$

$$V = S_{41}Z_i + S_{42}(dX_d/dt) + S_{43}D_o + S_{44}U \quad (3.25h)$$

where X_i is the independent energy (state) vector:

X_d is the dependent energy vector;

U is the system input vector;

V is the system output vector;

Z_i is the independent co-energy vector;

Z_d is the dependent co-energy vector;

D_i is the dissipation field input vector;

D_o is the dissipation field output vector

In fact, this method decomposes a system into the individual elements, and then connects them by the junction structure matrix. It solves not only linear problems, but also very general nonlinear problems.

3.7.2 Reciprocal system method

There has been an attempt to connect the bond graph modeling approach with singular perturbation theory (Dauphin-Tanguy, et al., 1985). It defines the notion of a reciprocal system and then applies the theory of perturbation. Thereby greater accuracy on the fast time scale behaviors of the system can be obtained. However, one must construct a reciprocal bond graph model. The analysis procedure can be illustrated as shown in Figure 3-1.

S1, S2, S3 and S4 are four different system models. The author claims that this method can be numerically implemented without difficulty, even if some matrix elements differ greatly in magnitude, because no matrix inversion is required.

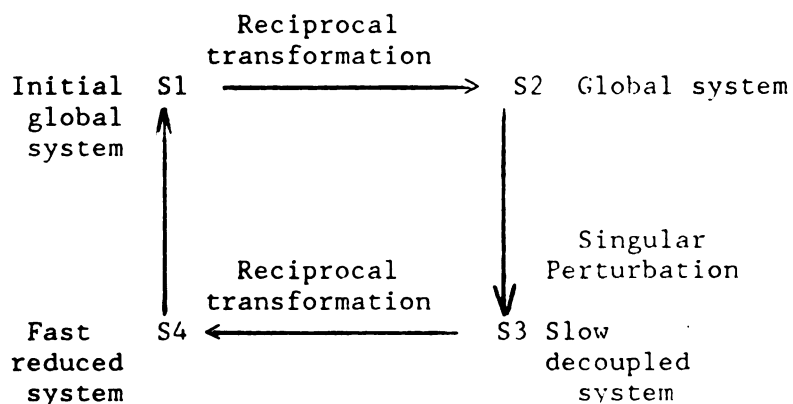


Figure 3-1 Reciprocal system method

3.7.3. Finite mode distributed system models

Finite mode distributed system models (Margolis, 1984) are extremely accurate in a chosen frequency range, while requiring only a fraction of the number of equations required by finite difference methods. The principal drawback is that the normal modes and frequencies must be obtained before the modeling process begins. This can be a tedious, if not impossible, task in itself. However, in many instances the actual system being modeled is composed of nearly uniform structural elements (such as beams, plates and membranes) for which the modes are easily obtained. Another problem usually associated with finite mode models is in the selection of appropriate boundary conditions for determining the original system modes.

The decoupled modal equations are obtained as

$$m_i \ddot{\eta}_i + k_i \dot{\eta}_i = \sum_{j=1}^J F_j W_i(x_j) \quad (3.26)$$

where $k_i = \omega_i^2 m_i$, m_i is the modal mass, ω_i is the frequencies for the unforced system, and $m_i = m \int_D W_i^2 dD$.

The actual system displacement is computed from

$$W(x,t) = \sum_{i=1}^{\infty} W_i(x) \eta_i(t) \quad (3.27)$$

The bond graph representation of Equation 3.26 is shown in Figure 3-2. Each of the retained modes is an I-C pair emanating from a common 1-junction. Each I element is a model mass m_i , while the corresponding

compliance, C , is an inverse modal stiffness or $1/m_i \omega_i^2$. The external forcing is properly represented by the S_e node associated with the discrete force $F_j(t)$. These forces are "felt" to each mode by multiplication with the proper mode shape evaluated at the location of the force. The TF moduli are nothing more than the mode shapes evaluated at the location of the respective forces, i.e., $TF_{11} = W_1(x_1)$, $TF_{12} = W_1(x_2)$, , $TF_{1k} = W_1(x_k)$, , $TF_{nk} = W_n(x_k)$. Thus the transformers (TFs) simply convert the actual force into modal forces while at the same time converting the modal velocities, $\dot{q}_i(t)$, into actual system velocities at the force location. In this fashion, the interactions between the continuum system and the lumped system are established.

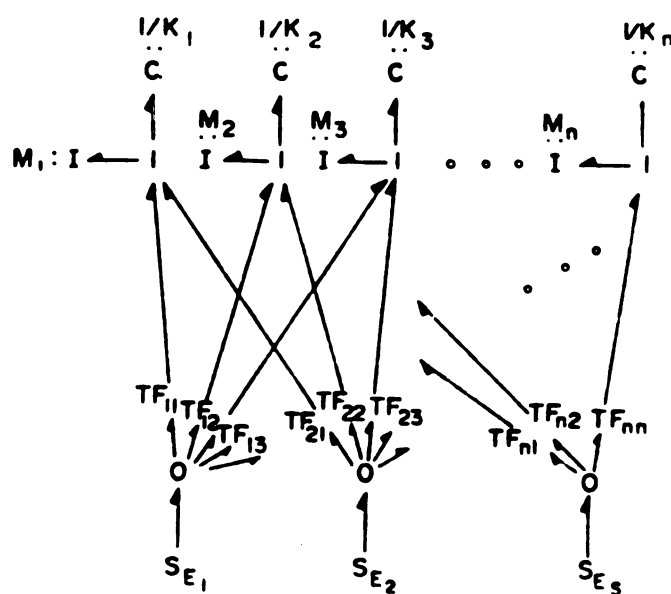


Figure 3-2 Finite modal bond graph model

Chapter 4

POWER POTRAITS OF DYNAMIC SYSTEM MODELS

All the model order reduction methods surveyed in Chapter 2 have a common aspect; namely, the starting point is the explicit system equations or transfer functions. For a physical system of large dimension, it is not easy work for an engineer to model it by a mathematical equation representation. Bond graph technology provides a potentially useful missing link between the physical system and its mathematical models. It is a computer-oriented modeling language and can be applied uniformly to many kinds of energy domains. In addition, it provides a clear picture of system topology. A new approach to model order reduction based on bond graphs is developed such that the model order reduction can be considered without first having to obtain the system equations. Let us next introduce the bond graph modeling technique.

4.1 Graphical Representations of Dynamic System Models

In addition to representing physical system models by explicit system equations, they may also be represented in graphical forms. Some examples are schematic diagrams, block diagrams, signal flow graphs, and bond graphs (Figure 4-1). Each of these representations has its own strengths and weaknesses. Schematic diagrams depict the configuration of physical systems in terms natural to particular physical domains, such as electrical and hydraulic circuits. The mathematical relations of the components are implied by the associated physical laws. The signal flow

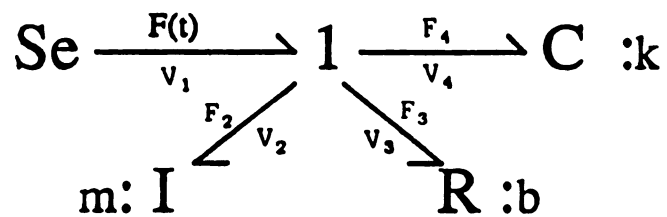
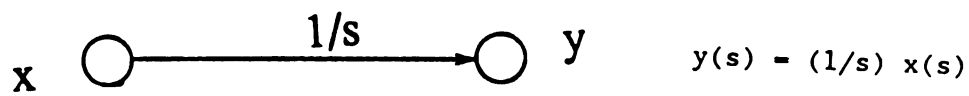
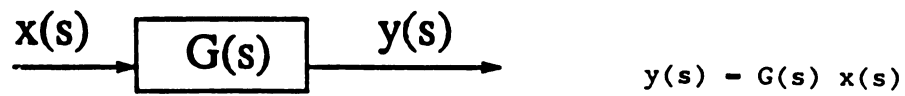
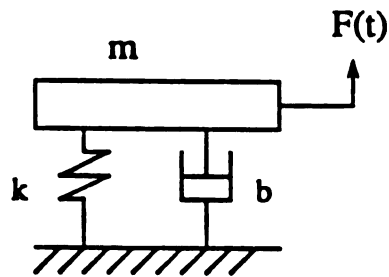


Figure 4-1 Graphical representations of mathematical modeling

graphs are similar in certain respects to block diagrams. Only are the block diagrams briefly discussed here.

Block Diagrams

Block diagram is a pictorial representation of the local cause and effect relationships among components of a model. Blocks may be aggregated or made more detailed as need arises. Block diagrams provide a convenient and useful representation for characterizing the functional relationships among the various components of control systems. The block diagram model stresses functional properties of modeled objects and their signal connections. The block diagram is a signal-based modeling method. Blocks of the model (i.e., nodes of the graph) are connected by directed lines which represent the direction of unilateral information or signal flow. Block diagram modeling has been widely used in control system design and simulation. In general, the block diagram is obtained by representing the particular mathematical equations, for example, the control laws. Block diagrams reveal signal relations clearly, but the power/energy aspects of a dynamic system are not readily accessible.

Bond Graphs

Power and energy attributes of models can be conveniently accessed by employing a power-based model representation, namely, bond graph model (Rosenberg and Karnopp, 1983). The bond graph is a pictorial modeling representation based on power coupling among components. A bond, the means of energy transfer between multiports, connects two ports. The bond graph is a power-based modeling method, since each bond connector contains a pair of power variables whose scalar product is the

power. A half-arrow on the bond represents the positive reference power direction. Using a rather small set of ideal elements, one can uniformly construct models of electrical, magnetic, mechanical, hydraulic, pneumatic, thermal, and other systems, or mixed systems. Standard techniques allow the models to be translated into a set of differential and algebraic equations by hand or by computer. The bond graph model depicts the physical effects considered by its modeler and their topological relations. It is easy in practice to modify the model structure to include additional effects. The bond graph technique has found many applications in engineering, biology, and even economics (Bos and Breedveld, 1985).

While bond graphs are excellent for modeling the 'plant' of a system, they are not well suited for modeling the controls of the plant. The signal communication aspect is better modeled by block diagrams. A need for a system model containing both bond graph and block diagram elements arises when the system being analyzed consists of subsystems which are best approached with different formalisms. The mixed bond graph/block diagram graph, developed by Zalewski and Rosenberg (1986) will be used in this study.

4.2 Uniform Performance Measure ---- Power

Power is often a neglected aspect of dynamic system response in system dynamics. We may use a very simple oscillation system as our illustration example. In Figure 4-2a a schematic diagram depicts the configuration and the linear parameters. By using Newton's law we may derive the equation of motion as follows

$$m \ddot{x} + b \dot{x} + k x = F(t) \quad (4.1)$$

where a superdot denotes a time derivative.

Figure 4-2b is the equivalent block diagram of the system. You may well ask: "where is the power?" Neither the schematic diagram nor the block diagram can give you the answer directly. However, aspects of the power response can be accessed and displayed clearly in its bond graph, Figure 4-2c. Each bond in the graph contains a pair of power variables, effort $e(t)$ (force) and flow $f(t)$ (velocity), whose product is power. In this example, the bond graph gives the following constitutive and connective equations:

$$\dot{p}_2 = F_2 \quad (4.2)$$

$$V_2 = m^{-1} p_2 \quad (4.3)$$

$$\dot{x}_4 = V_2 \quad (4.4)$$

$$F_4 = kx_4 \quad (4.5)$$

$$F_3 = bV_3 \quad (4.6)$$

$$F_2 = F_1 - F_3 - F_4 \quad (4.7)$$

$$V_1 = V_2 = V_3 = V_4 \quad (4.8)$$

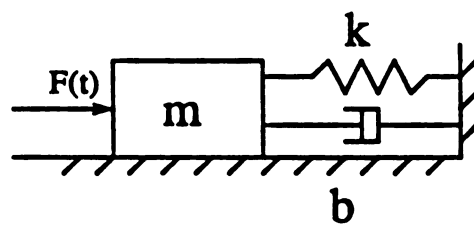
where F_i is force on bond i ,

V_i is velocity on bond i ,

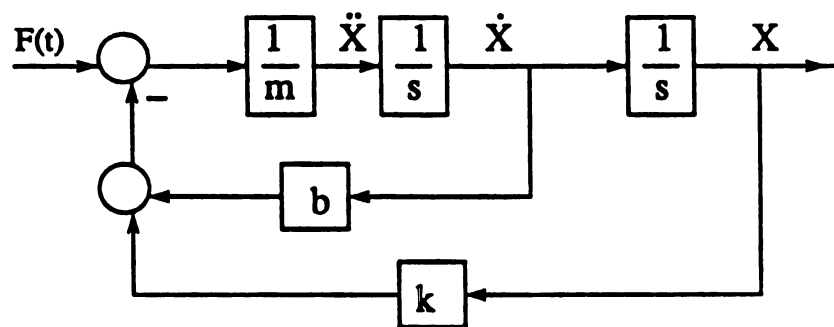
p_i is momentum on bond i , and

x_i is displacement on bond i .

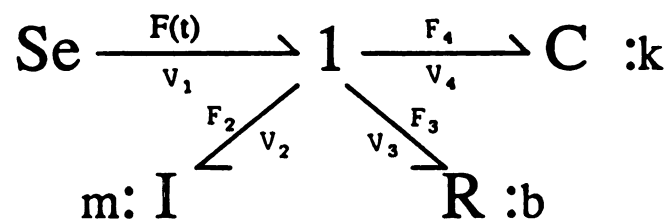
Also associated with each bond in this bond graph is the power. W_i , denoting the power on bond i , is the product of F_i and V_i . Also the net energy transfer, E_i can be found from $\int_t W_i d\tau$.



(a)



(b)



(c)

Figure 4-2 Powers in system models

(a) schematic diagram

(b) block diagram

(c) bond graph

Although the power variables (effort $e(t)$ and flow $f(t)$) in different energy domains have different physical meanings and units, their product in every domain is power. As long as the efforts and flows have appropriate units (e.g., in the SI unit system), their power products are the same (e.g., watt). Table 4-1 summarizes the power variables of the bond graph in several energy domains. We can see that power is a uniquely uniform variable in a multi-domain system. Some properties of physical systems may be investigated efficiently by using such a uniform basis.

Table 4-1 Powers in various energy domains

Domain	Effort, $e(t)$	Flow, $f(t)$	Power, $P(t)$
Mechanical Translation	Force, $F(t)$	Velocity, $V(t)$	$F(t)V(t)$
Mechanical Rotation	Torque, $\tau(t)$	Angular Vel., $\omega(t)$	$\tau(t)\omega(t)$
Electrical Circuit	Voltage, $V(t)$	Current, $I(t)$	$V(t)I(t)$
Hydraulic Circuit	Pressure $P(t)$	Volume Flow Rate $Q(t)$	$P(t)Q(t)$
Magnetic Circuit	Mmf $M(t)$	Flux Rate $\dot{\Phi}(t)$	$M(t) \dot{\Phi}(t)$
Thermal	Temperature $T(t)$	Entropy Rate $\dot{S}(t)$	$T(t) \dot{S}(t)$
Heat Transfer	Temperature $T(t)$	Heat Flow Rate $Q(t)$	$T(t)Q(t)$ *

* Not a true power.

It is natural at this point to raise the question, "How can we exploit the knowledge about powers and energy transfer in a bond graph model for better understanding of the system dynamics?" Answers may make a valuable tool for model order reduction. The first step is to make the power information readily available and accessible.

4.3. Power/energy Visualization

4.3.1 Computation of power variables

Referring the diagram in Figure 2-2 and 2-3 of Chapter 2, the source field, energizer storage field, dissipation field, and transducer field are coupled by a power-conserving connective multiport represented by the junction structure. The Paynter junction structure consists of 0 and 1 junction elements, (It is named after H. M. Paynter, the inventor of the bond graph.) and it is invariant. The modulated junction structure is the collection of the transformers and gyrators, which may have varying moduli. The key vectors of each field are identified. The system equations are defined for each field as follows:

$$\text{Source field} \quad U = \Phi_s(t) \quad (4.9)$$

$$\text{Storage field} \quad Z_i = \Phi_i(X_i, X_d) \quad (4.10)$$

$$Z_d = \Phi_d(X_i, X_d) \quad (4.11)$$

$$\text{Dissipation field} \quad D_o = \Phi_L(D_i) \quad (4.12)$$

Transducer coupling structure

$$T_o = \Phi_t(X_i, t) * T_i \quad (4.13)$$

The invariant connective structure is represented by the junction matrix and the associated vectors as follows

$$V_o = S * V_i \quad (4.14)$$

$$V_i = [Z_i, X_d, D_o, U, T_o]^t$$

$$V_o = [X_i, Z_d, D_i, V, T_i]^t$$

where V_i is the input vector to the junction field and V_o is the output vector of the junction field.

Furthermore, the variables of bonds incident to only 0 and 1 junctions are collected in a single vector, Y , which includes both effort variables and flow variables. The connective relations in the Paynter junction structure are defined by the square matrix P which consists only of the integers 0, 1, and -1. The P matrix can be decomposed such that

$$V_o = P_{oi} * V_i + P_{oy} * Y \quad (4.15a)$$

$$Y = P_{yi} * V_i + P_{yy} * Y \quad (4.15b)$$

Since the coefficients of P are all constants, provided the required inverse exists, we may find a solution for Y from

$$Y = ((I - P_{yy})^{-1} P_{yi}) * V_i \quad (4.16)$$

Using (4.16) in (4.15a) we can solve the output vector for the Paynter junction structure as

$$V_o = (P_{oi} + P_{oy}((I - P_{yy})^{-1}P_{yi}))^*V_i \quad (4.17)$$

Thus far, we are able to obtain every pair of power variables in a bond graph model. Therefore we can evaluate the power response on bond i simply by multiplying the proper pair of the two power variables; namely,

$$W_i(t) = e_i(t)*f_i(t) \quad (4.18)$$

The powers in a bond graph may be classified into two major types, namely, external and internal. The external powers are those on the external bonds which connect physical nodes (Se, Sf, C, I, R), while the internal powers are those on the internal bonds which are incident only to 0 or 1 junctions. The power and energy flow on the external bonds are defined as follows:

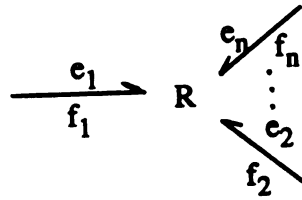
Source elements

← Se (Sf)

$$W = e*f \quad (4.19a)$$

$$\bar{w} = \int_{t_1}^{t_2} W \, d\tau \quad (4.19b)$$

The net energy transfer on the bond represents the energy supplied to the system during the time period from t_1 to t_2 .

Resistance elements

For passive resistors, the power W is always positive, i.e.

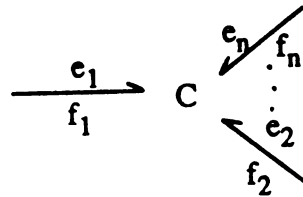
$$W_i = e_i * f_i \geq 0$$

The energy transfer on the bond represents the energy consumption during the time period from t_1 to t_2 .

$$T_i = \int_{t_1}^{t_2} e_i * f_i \, d\tau \quad (4.20a)$$

The total energy absorbed is

$$E = \sum_{i=1}^n T_i \quad (4.20b)$$

Capacitance elements

For springs and other types of capacitors, the power represents the time rate of change of energy storage. If $W > 0$ the capacitor receives

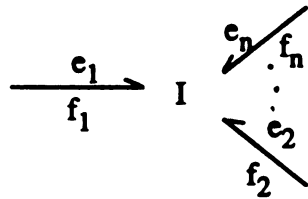
energy, while for $W < 0$ the capacitor releases energy. The net energy transfer on a bond is computed by

$$T_i = \int_{t_1}^{t_2} W_i \, dr = \int_{t_1}^{t_2} e_i * f_i \, dr \quad (4.21a)$$

The energy stored is given by

$$E = \sum_{i=1}^n T_i + E_0 \quad (4.21b)$$

Inertance elements



For masses and other types of inertial elements, the power W is the time rate of change of kinetic energy. If $W > 0$ the inertial element receives energy, while for $W < 0$ the inertial element loses kinetic energy. The net energy transfer on a bond is computed by

$$T_i = \int_{t_1}^{t_2} W_i \, dr = \int_{t_1}^{t_2} e_i * f_i \, dr \quad (4.22a)$$

The energy stored is given by

$$E = \sum_{i=1}^n T_i + E_0 \quad (4.22b)$$

Internal bonds

The power on an internal bond is the product of its power variables

$$W_i = e_i * f_i$$

4.3.2 Power measures

In a dynamic system the power on a bond is a function of time that is positive when the power is in the half-arrow direction, or negative, if it opposes the half arrow. The power history can be displayed in a usual way to assist system analysis. Figure 4-3 is the power responses versus time of the example system of Figure 4-2 for input $F(t) = 1.0$, where $W.B1$ is the power on bond B1, and so on. Note that $W.B3$ is always greater than zero since it is dissipated from this system by the damper.

The power conservation $\sum_{i=1}^4 W.Bi = 0$ is true at all time.

Power responses versus time show how the powers associated with different types of physical effects in different energy domains vary. However, for a large scale dynamic system its bond graph may contain tens or even hundreds of external bonds and their associated physical effects, plus many internal bonds. While it is possible to display every power response history for all bonds in a plot, it is very difficult to abstract useful information from the vast amount of data. An alternative way is to display aspects of the power response on the bond graph itself. Since the power can be evaluated over a time period, an averaged power over that period may reflect the intensity of the local interaction within the global system. The bond provides the perfect

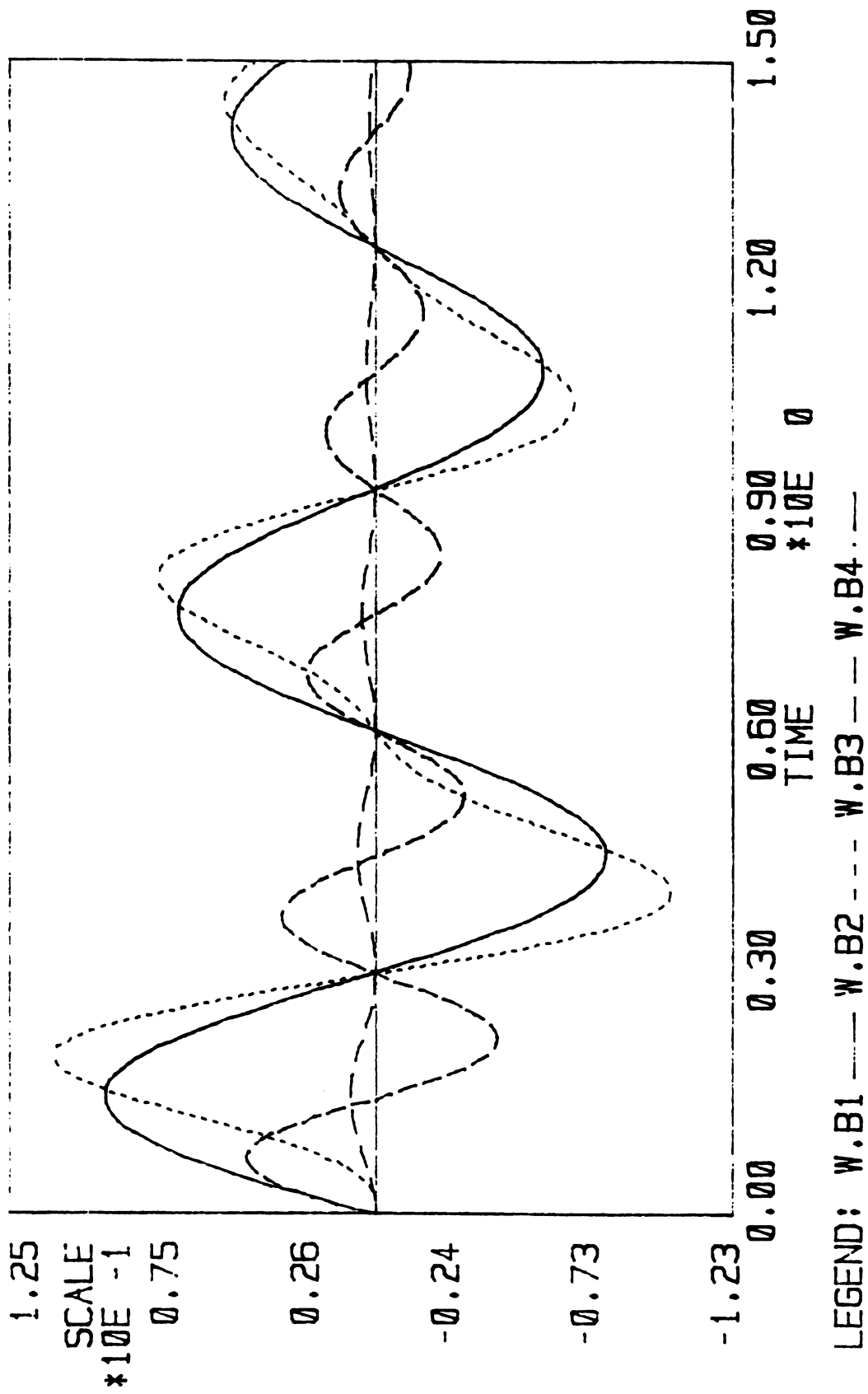


Figure 4-3 Power responses of the system in Figure 4-2

place to display this intensity quantitatively by color, width of the bond, or some other means.

For the purpose of possible decomposition and simplification of the system model, a statistical measure is introduced as below. The magnitude of the power reflects the strength of the interaction on a bond. For a suitable time period the magnitudes of averaged powers on all bonds are computed and compared to find the relative interaction strengths. Some possible power measures are:

$$\text{arithmetic mean} \quad \bar{W} = (1/T) \int_T W \, d\tau ; \quad (4.23a)$$

$$\text{absolute mean} \quad \bar{W} = (1/T) \int_T |W| \, d\tau ; \quad (4.23b)$$

$$\text{and root mean square} \quad \bar{W} = [(1/T) \int_T W^2 \, d\tau]^{1/2} \quad (4.23c)$$

For ease of implementation the approximate measures corresponding to the above definitions are employed:

$$\text{arithmetic mean} \quad \bar{W} = (1/K) \sum_{k=1}^K W_k ; \quad (4.24a)$$

$$\text{absolute mean} \quad \bar{W} = (1/K) \sum_{k=1}^K |W_k| ; \quad (4.24b)$$

$$\text{and root mean square} \quad \bar{W} = [(1/K) \sum_{k=1}^K W_k^2]^{1/2} ; \quad (4.24c)$$

where K is the total number of data stored and W_k is the power at the step k . The range for k can be subinterval within the computed range. Each operation may be used to make the best display in different cases. Powers and energy transfers on all bonds can be computed. The energy of the external nodes may also be computed.

4.3.3 Scaling and color-coding

Color graphics techniques are being used in the practical worlds of science and industries as well, to help manage and interpret the vast amounts of scientific and technical data. Graphics reveal trends and relationships that would otherwise remain buried under mountains of numerical detail. With the aid of computer, graphic information is generated, analyzed, and displayed for the power properties in a bond graph model of a dynamic system. A color graphical display of the power strength can be shown on a computer screen, which gives user very clear picture of the power distribution in the bond graph.

In the power display the color of a bond is determined by a certain scale and classified into a finite groups. There are three scales available in the computer implementation. They are linear, logarithmic, and rank-ordered. All bonds are sorted into six groups, each of which is assigned a color from red (highest power) to blue (lowest power) in the descending order of magnitude.

4.4 Power Method for Model order Reduction and Model Simplification

Since any physical system is energy related, in addition to the generalized momenta and displacements and the generalized efforts and flows, the power and energy should also provide insight into the system dynamics. As discussed above the power on a bond is the product of its two power variables, effort e and flow f . The ideal elements are connected by the bonds which are the energy pipes between the connected elements or subsystems. The time integration of the power is the energy flow through the bond over the time period. Power can be treated as the

uniform indication of the interaction between two sides of the bond. It can be analogue to an electrical system in which the current through a line and the voltage across the line are important attributes, but their product, power, also provides additional knowledge about the system.

The basic idea behind the notion of power analysis is that the importance (weight) of a physical effect or a subsystem to the global system dynamics may be evaluated by the strength of its interaction with the rest of the system. The averaged power at the bond could be used as a measure of the strength, therefore, the significance of a physical effect or a subsystem can be revealed.

The simplification of a dynamic model can be achieved in two stages. The first is at the subsystem level and the second is at the component level. At the subsystem level we seek to decompose a bond graph model into a number of subsystems, based on physical insight, linear/nonlinear separation, and the topology of the junction structure. At the component level, the importance of each physical effect is evaluated based on the power interaction with the whole system.

4.4.1 Simplification of physical effects

A power display on a bond graph as shown in Figure 4-4 gives a clear picture of the power distribution with respect to the physical effects. A set of the weakest physical effects can be identified immediately from the powers on the external bonds. Since the power on an external bond reflects the strength of the interaction with the rest of system, a measure of the power magnitude may suggest the importance of the associated effect. Weak physical effects can be removed from the

system on a trial basis. To guide the removal of the weakest physical effects, some criterion must be set according to the nature of the effects, the power metrics, and accuracy of the approximation. One potentially useful measure is the effect index, EI, defined as follows:

$$EI_i = (\hat{W}_i / \hat{W}_{\max}) * 100 \% \quad (4.25)$$

where $\hat{W}_i$ is the absolute mean (or root mean square) power of the i -th bond, and $\hat{W}_{\max}$ is the maximum absolute mean (or root mean square) power among the all bonds. The use of the effect index will be illustrated through an example in Section 4.5.

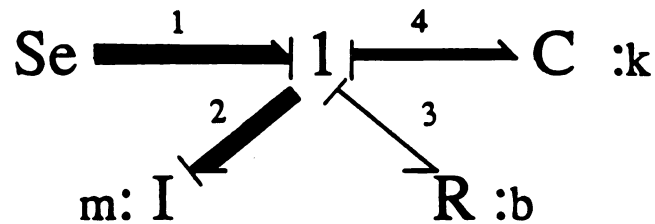


Figure 4-4 Powers on external bonds

4.4.2 Decomposition and simplification of subsystems

Elimination of unimportant physical effects may yield a reduced order model if the ignored effects include the capacitance or inductance. A model order reduction in the form of decoupling also can be achieved by decomposing the system into subsystems by removing the least significant subsystems whose connecting bonds have the lowest level powers. In this case, a cut set of bonds with the low power levels indicate the weak connections between the subsystems and the "mother" system, even though the bonds inside of the subsystems may have high level powers. For example, assuming that in Figure 4-5, bonds 1, 2, and up to k are internal and they are cut bonds where the system is decomposed into k subsystems S_1 , S_2 , and up to S_k . By examining the powers on these bonds we found that one of them, say bond 1, has very small power compared with the rest of bonds. It shows that subsystem S_1 has weak connection to the other subsystems: therefore we can simplify the system by removing subsystem S_1 with a low risk of removing significant effect. The resulting computations will involve separated subsystems.

Similar to the effect index, the connection index, CI, is defined as follows:

$$CI_i = (\bar{W}_i / \sum_j \bar{W}_j) * 100 \% \quad < 1.0 \quad (4.26)$$

where $\bar{W}_i$ is the root mean square (or absolute mean) power of the cut bond of the i-th subsystem, $\bar{W}_j$'s are the power on all the cut bonds connected to the junction structure.

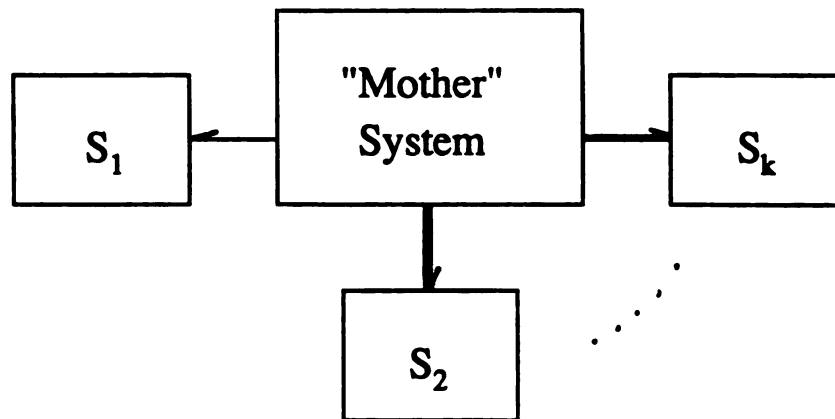


Figure 4-5 Powers on cut bonds

4.4.3 Procedure for model simplification

By using the above suggested effect index and connection index a general procedure to simplify the system model in a bond graph framework has been developed. The procedure is described as below:

1. Conduct a simulation under nominal conditions on inputs and parameters.
2. Identify and remove insignificant subsystems:
 - a. By applying the connection index criterion, identify the weakest cut bonds from all cut bonds, thereby the most insignificant subsystems:

b. Remove the most insignificant subsystems.

3. Eliminate the unimportant physical effects:

a. By applying the effect index criterion, identify the weakest external bonds, thereby the most unimportant physical effects;

b. Remove the most unimportant physical effects

The developed method would be best illustrated by some examples in the following section.

4.5. Examples

4.5.1. A Radar pedestal unit

For illustration, let us model a radar pedestal unit and its control. A sketch of the radar pedestal unit is shown in Figure 4-6. The purpose of the drive motor and control system is to set the angular position of the pedestal about its vertical axis as desired. The system graph for pedestal position control is shown in Figure 4-7. The system graph includes a bond graph part for the unit and a block diagram portion for the feedback control.

The bond graph part models a DC drive motor connected through a shaft and gears to a pedestal unit. The input voltage to the drive motor is supplied by node SEE. The dominant effects in the electrical part of the motor are the inductance (IE) and the resistance (RE). The field current (signal I) is converted to a torque with no back-voltage effect.

IM and RM are the motor inertia and friction. The shaft has torsional stiffness (CS). The gears are modeled as an ideal transformer (TF). The pedestal load is composed of a rotational inertia (IP) and a friction effect (RP). Two types of power bonds are used, the electrical set (E1, E2, E3) and the mechanical rotation sets (M1, M2, M3 for the motor), (S1, S2, S3 for the shaft), and (P1, P2, P3 for the pedestal). The bond graph part, starting with the motor input SEE, can be used to investigate the open-loop characteristics of the drive system and load.

The position feedback control part of the model includes a transfer function, $1/s$, which integrates the velocity of the pedestal, w , to get its position θ . Node SUM compares the position with the desired value, REF, the error CV is sent to modulate the input voltage of the drive motor.

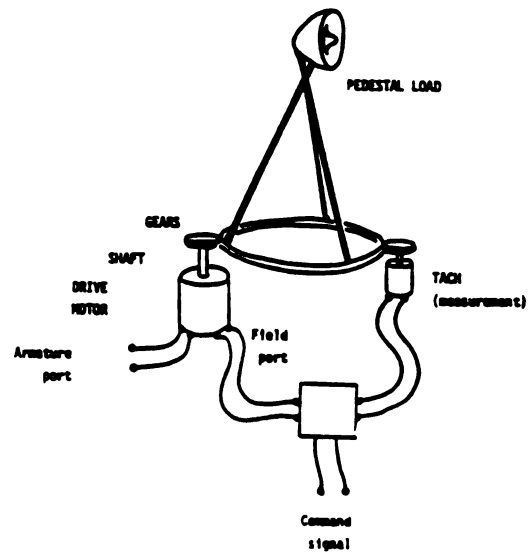


Figure 4-6 A radar pedestal unit

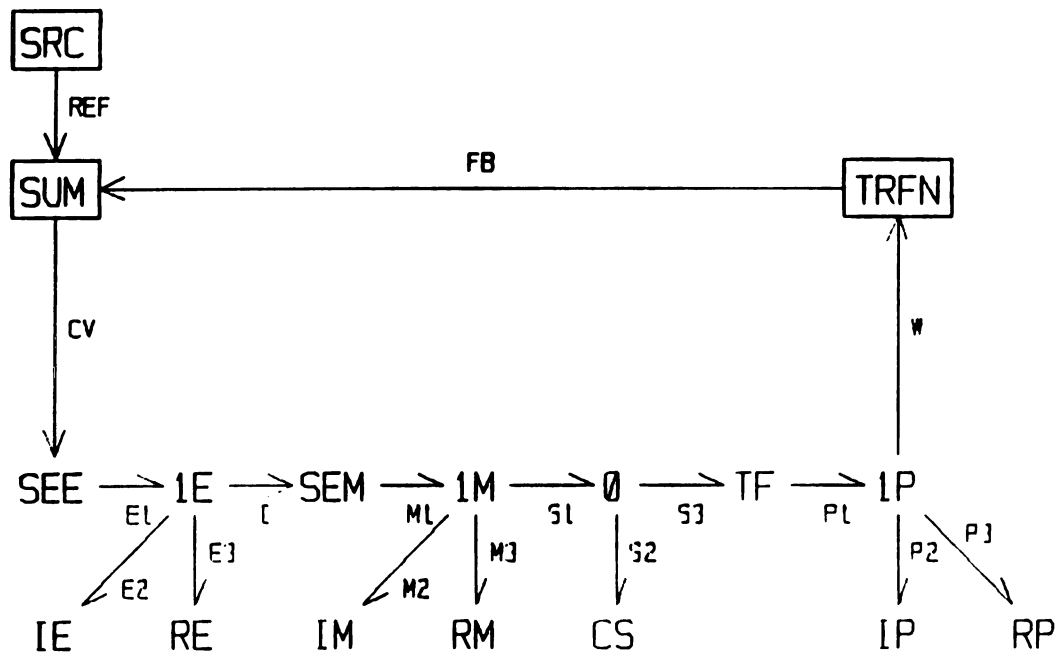


Figure 4-7 Bond graph for the radar pedestal unit

The state equations for the radar pedestal unit system can be derived as follows:

$$\dot{p}_1 = -(r_1/I_1)p_1 + K_e(u - \theta) \quad (4.27a)$$

$$\dot{p}_2 = -(r_2/I_2)p_2 - kq + (K_m/I_1)p_1 \quad (4.27b)$$

$$\dot{q} = (1/I_2)p_2 - (m/I_3)p_3 \quad (4.27c)$$

$$\dot{p}_3 = mkq - (r_3/I_3)p_3 \quad (4.27d)$$

$$\dot{\theta} = (1/I_3)p_3 \quad (4.27e)$$

The system matrix A and the input matrix B are as follows:

$$A = \begin{bmatrix} -r_1/I_1 & 0 & 0 & 0 & -K_e \\ K_m/I_1 & -r_2/I_2 & -k & 0 & 0 \\ 0 & 1/I_2 & 0 & -m/I_3 & 0 \\ 0 & 0 & mk & -r_3/I_3 & 0 \\ 0 & 0 & 0 & 1/I_3 & 0 \end{bmatrix}$$

$$B = [K_e \ 0 \ 0 \ 0 \ 0]^T$$

The parameters for this system are given in Table 4-2.

Table 4-2 Parameters of pedestal model

$I_E = 0.1$ henrys	$r_E = 5.0$ ohms
$I_M = 0.25$ kg.m ²	$r_M = 0.3$ N.s/rad
$k_S = 100.0$ N/rad	$I_P = 320.0$ kg.m ²
$r_P = 50.0$ N.s/rad	$K_e = 1.0$
$M1 = 20.0 * I$	$TF = 30.0$

Therefore the matrix A and B are evaluated as:

$$A = \begin{bmatrix} -50 & 0 & 0 & 0 & -1 \\ 200 & -1.2 & -100 & 0 & 0 \\ 0 & 4 & 0 & -0.09375 & 0 \\ 0 & 0 & 3000 & -0.15625 & 0 \\ 0 & 0 & 0 & 0.003125 & 0 \end{bmatrix}$$

$$B = [1 \quad 0 \quad 0 \quad 0 \quad 0]^t$$

The simplification by removing the weakest physical effects may be seen clearly from the power distribution of the radar pedestal unit. An RMS measure of the power on each bond has been computed and the resulting list of values has been rank-ordered. The abstracted power response of the system is listed in Table 4-3. The column at the right lists the RMS power value as a percentage of the maximum value. Figure

Table 4-3 Effect indices

ORDER	BOND	RMS POWER	EI %
1	M1	0.1059E+02	100.000
2	M3	0.8120E+01	76.642
3	S3	0.4030E+01	38.041
4	P1	0.4030E+01	38.041
5	S1	0.4023E+01	37.969
6	P2	0.3739E+01	35.287
7	M2	0.3239E+01	30.575
8	P3	0.1506E+01	14.210
9	S2	0.2620E+00	2.473
10	E3	0.8665E-01	0.818
11	E1	0.8662E-01	0.818
12	E2	0.1711E-01	0.016

Figure 4-8 shows the root mean square powers displayed in six colors. We found that bonds E1, E2, and E3 have the lowest power levels. However, these are the powers in the controlled field of the motor. The major power supply of the system is from the armature. The connector I separates the two parts. In the main part bond S2 has the smallest effect index with 2.473%. The physical effect associated with bond S2 is the shaft compliance. It is expected that the shaft transmits large torque with small rate of change of the torsional displacement under a high oscillation frequency. It shows that the compliance of the shaft is not significant with respect to the general system dynamics and suggests that the ignoring the element CS would not degrade the model. The simplified model is shown in Figure 4-9. Since the removal of node CS causes derivative causality on one of the bonds M2 and P2, the two inertance IM and IP are combined into a single equivalent inertance IEQ. The order of the simplified model is reduced by two. The eigenvalues of these two models are listed in Table 4-4 and their responses are plotted in Figure 4-10. It is not surprising that the modes with high frequency are truncated and the responses are almost identical except for the ripples on the curve of the angular velocity of the original system.

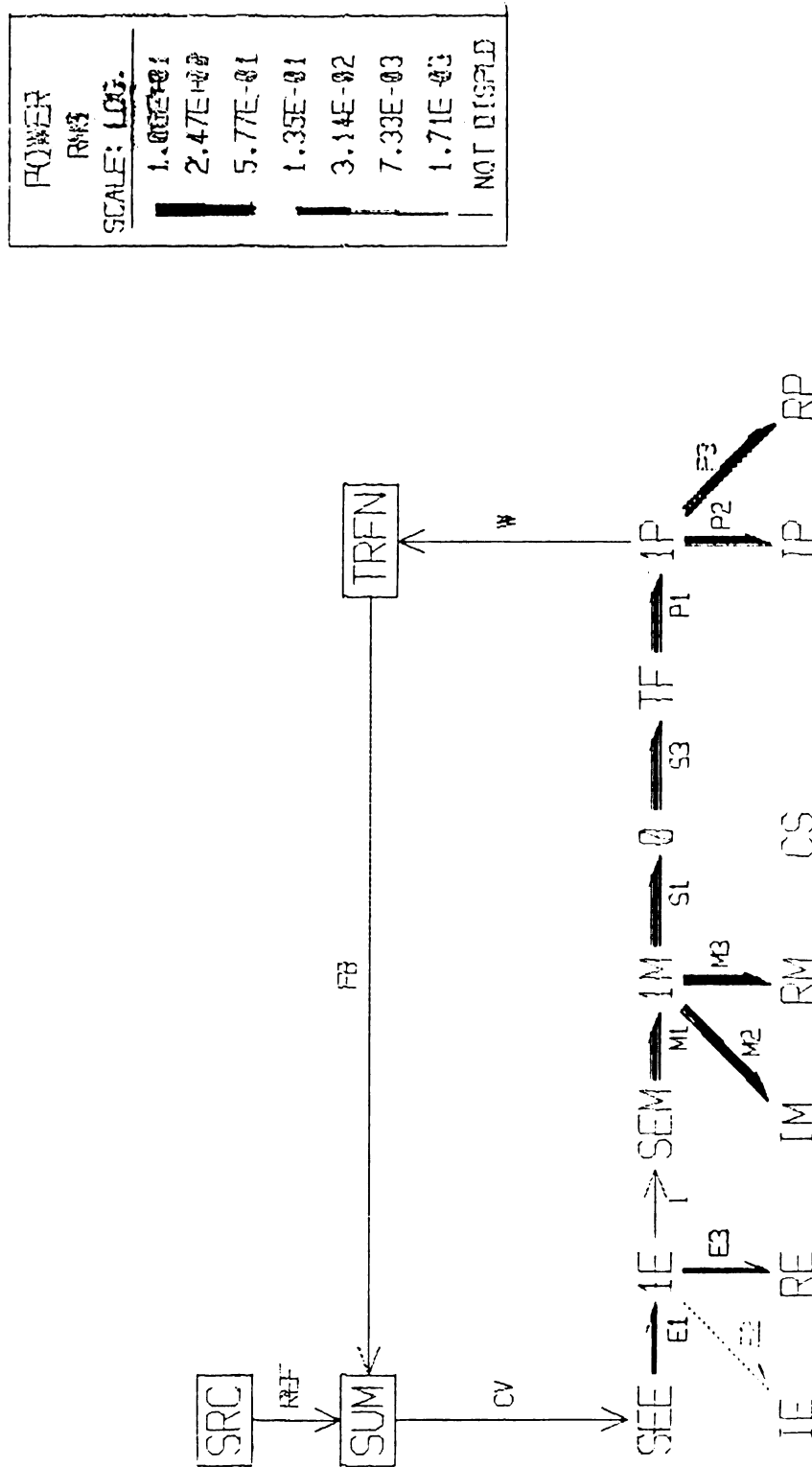


Figure 4-8 Power display on the bond graph of the pedestal

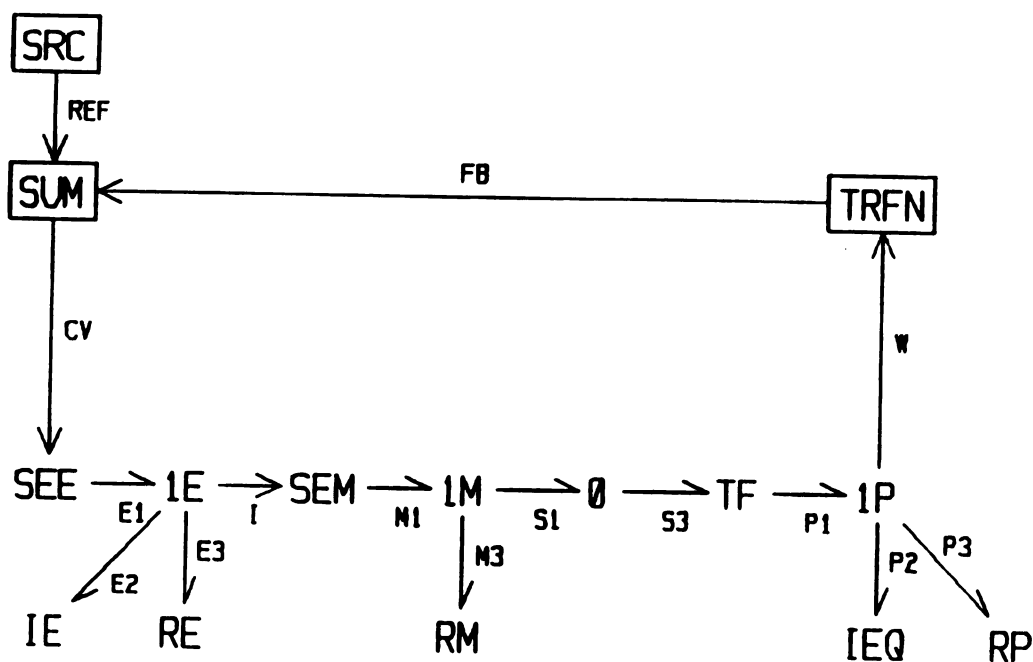


Figure 4-9 Reduced-order model of the pedestal

Table 4-4. Eigenvalues of the original and reduced models

Original	Reduced
-3.86210E-01 ± j 2.60896E+01	
-2.91432E-01 ± j 3.67950E-01	-2.91340E-01 ± j 3.67800E-01
-5.00010E+01	-5.00045E+01

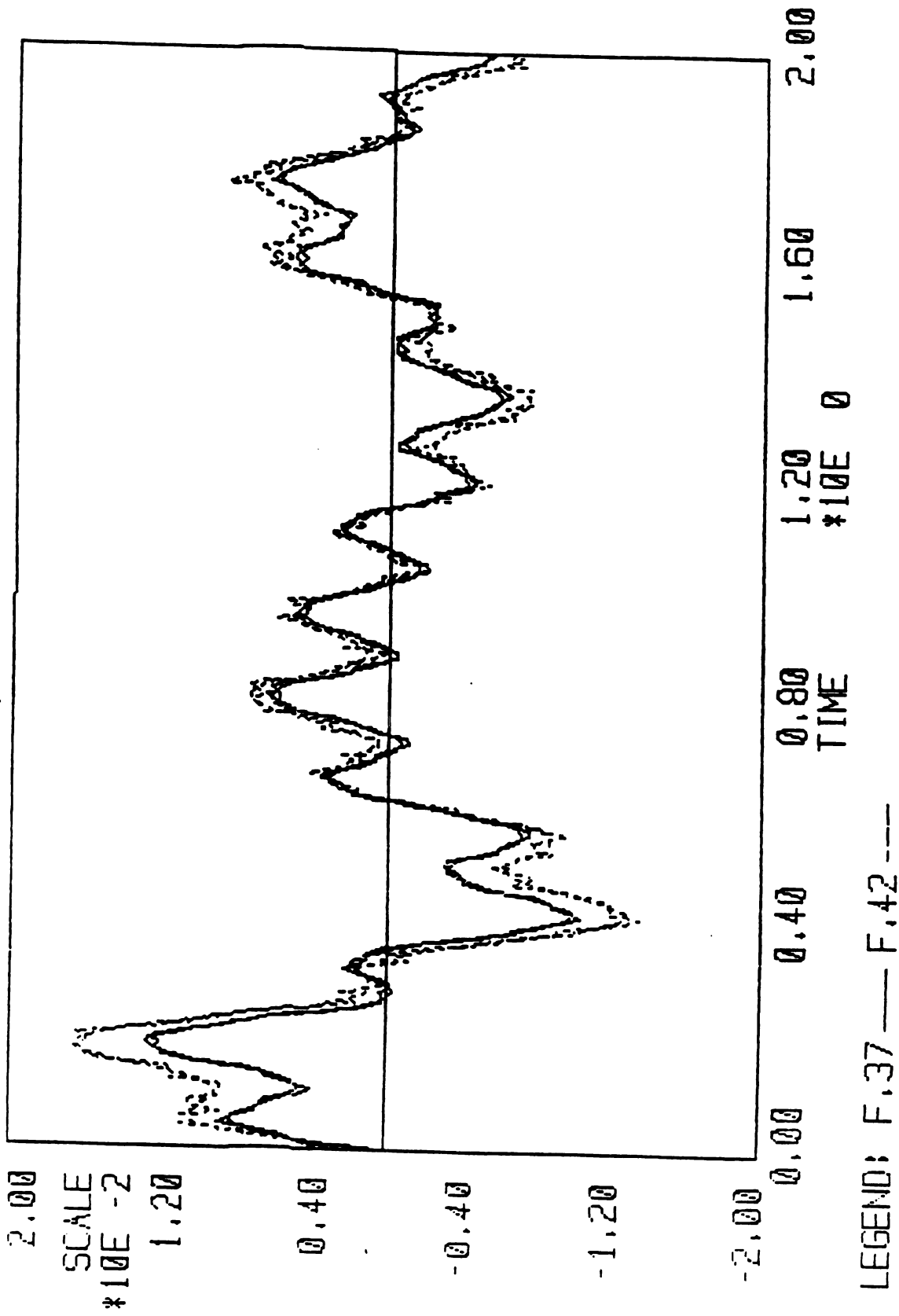


Figure 4-10 Responses of the original and reduced models

4.5.2 An Euler-Bernoulli beam with a vibratory load

As a second example consider a pinned beam with a vibratory load attached at position x_1 and a driving force $F(t)$ acting at position x_2 , as shown in Figure 4-11. The beam acts as a coupling element between a driving force and the load. Following the modal bond graph approach of Karnopp and Rosenberg (1968) and Margolis and Tabrizi (1984) the continuous beam is modeled by a modal bond graph with five modes retained (See Figure 4-12). Mode 1 is represented by $C1, I1$; mode 2 is represented by $C2, I2$; The modal frequencies of the beam are given by

$$\omega_i = (i \pi)^2 \sqrt{EI/\rho AL^4} \quad i = 1, 2, \dots \quad (4.28)$$

where ω_i is the i -th modal frequency; E is the Young's modulus; I is the area moment of inertia; ρ is the density of the beam material; and L is the length of the beam. In this example, we assume $\sqrt{EI/\rho AL^4} = 1.0$, so that $\omega_1 = \pi^2 = 9.87$, $\omega_2 = (2\pi)^2 = 39.48$, $\omega_3 = (3\pi)^2 = 88.83$, $\omega_4 = (4\pi)^2 = 157.91$, $\omega_5 = (5\pi)^2 = 246.74$. The model mass m_i is determined as follows

$$\begin{aligned} m_i &= \int_0^L \rho A Y_i^2(x) dx = \int_0^L \rho A \sin^2(i\pi x/L) dx \\ &= \int_0^1 \rho AL \sin^2(i\pi x/L) d(x/L) \\ &= \rho AL [1 - \cos(i\pi x/L)]/2 d(x/L) \\ &= \rho LA/2 = M/2 \end{aligned} \quad (4.29)$$

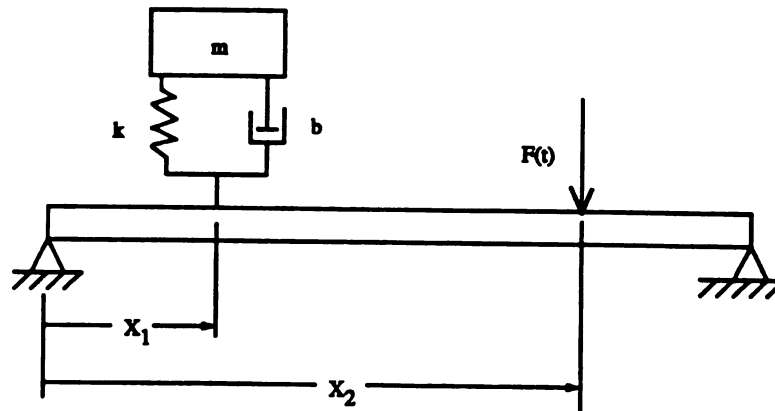


Figure 4-11 A beam-coupled system

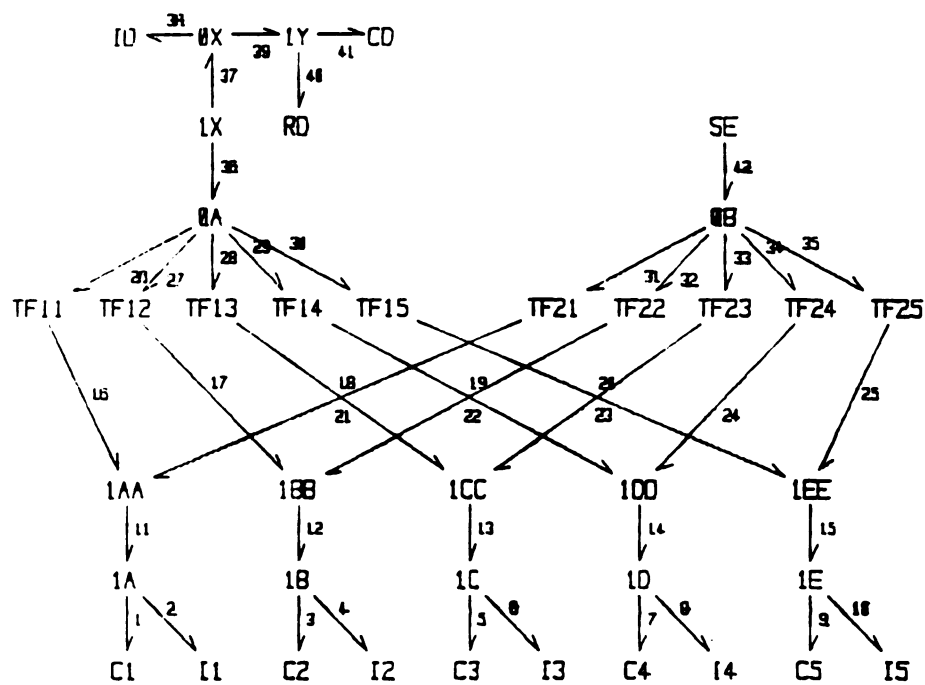


Figure 4-12 Modal bond graph model for the beam-coupled system

where M is the mass of the beam. Let $M = 10.0$ kg; then $m_1 = m_2 = m_3 = m_4 = m_5 = 5.0$ kg. The modal stiffnesses are computed by

$$\begin{aligned} k_1 &= m_1 \omega_1^2 = 487.08 \text{ N/m} \\ k_2 &= m_2 \omega_2^2 = 7793.4 \text{ N/m}, \\ k_3 &= m_3 \omega_3^2 = 39453.8 \text{ N/m}, \\ k_4 &= m_4 \omega_4^2 = 124677.8 \text{ N/m}, \\ k_5 &= m_5 \omega_5^2 = 304403.1 \text{ N/m}. \end{aligned}$$

The parameter values for the lumped mass and spring are chosen to be $m = 2.0$ kg and $k = 200.0$ N/m. Thus the load natural frequency is $\omega = 10.0$. For a damping ratio 0.1, we choose $b = 4.0$ Ns/m.

The mode shapes at position x_j are determined by

$$Y_j^i = \sin(j\pi x_i/L) \quad i = 1, 2, \dots; \quad j = 1, 2, \dots \quad (4.30)$$

The mode shapes at the positions of the prescribed force $F(t)$ and the load are calculated from Equation 4.30 when $x_1 = L/5$ and $x_2 = 2L/3$, where L is the beam length. They are

$$\begin{aligned} Y_1 &= [0.5878, 0.9511, 0.9511, 0.5878, 0.0]^T, \\ Y_2 &= [0.866, -0.866, 0.0, 0.866, -0.866]^T \end{aligned}$$

The RMS power distribution for a unit step force input (E.42) in the bond graph is given in Table 4-5. The data are coded on the bond graph in Figure 4-13. Bonds 30 and 20 have zero power because the load is located at a node of mode 5 ($x_1 = L/5$). Bonds 33 and 23 have zero power because the input force is located at a node of mode 3 (i.e., $x_2 = 2L/3$). The lower part of the bond graph is the modal representation of

Table 4-5 The power listing of five-mode model

ORDER	BOND	RMS POWER	EI %
1	1	0.6239E-02	100.000
2	42	0.6181E-02	99.079
3	11	0.5808E-02	93.098
4	21	0.5783E-02	92.692
5	31	0.5783E-02	92.692
6	2	0.2155E-02	34.535
7	3	0.2071E-02	33.188
8	12	0.1931E-02	30.956
9	22	0.1892E-02	30.327
10	32	0.1892E-02	30.327
11	38	0.1587E-02	25.443
12	39	0.1491E-02	23.900
13	41	0.1403E-02	22.487
14	36	0.8567E-03	13.731
15	37	0.8567E-03	13.731
16	4	0.7426E-03	11.903
17	16	0.7200E-03	11.541
18	26	0.7200E-03	11.541
19	7	0.6410E-03	10.275
20	14	0.5910E-03	9.474
21	34	0.5850E-03	9.378
22	24	0.5850E-03	9.378
23	40	0.4760E-03	7.630
24	9	0.4250E-03	6.813
25	17	0.3908E-03	6.264
26	27	0.3908E-03	6.264
27	15	0.3871E-03	6.205
28	25	0.3871E-03	6.205
29	35	0.3871E-03	6.205
30	8	0.2477E-03	3.970
31	10	0.1760E-03	2.821
32	29	0.7394E-04	1.185
33	19	0.7394E-04	1.185
34	5	0.7855E-05	0.126
35	13	0.7753E-05	0.124
36	18	0.7753E-05	0.124
37	28	0.7753E-05	0.124
38	6	0.3685E-06	0.006
39	20	0.0000E+00	0.000
40	23	0.0000E+00	0.000
41	30	0.0000E+00	0.000
42	33	0.0000E+00	0.000

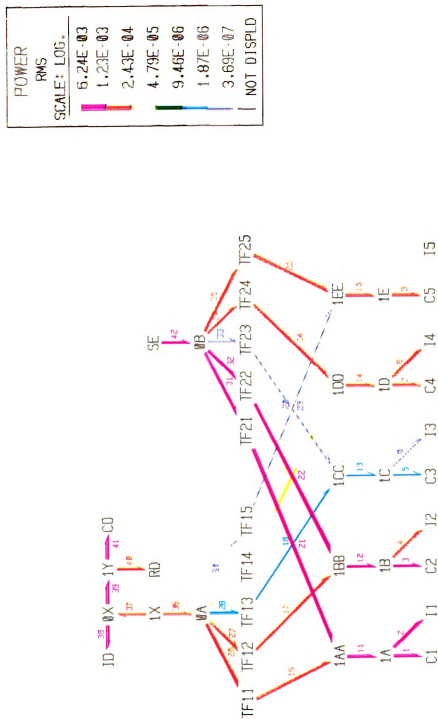


Figure 4-13 Power distribution in the modal bond graph model

POWER
RMS
SCALE : 1.05.

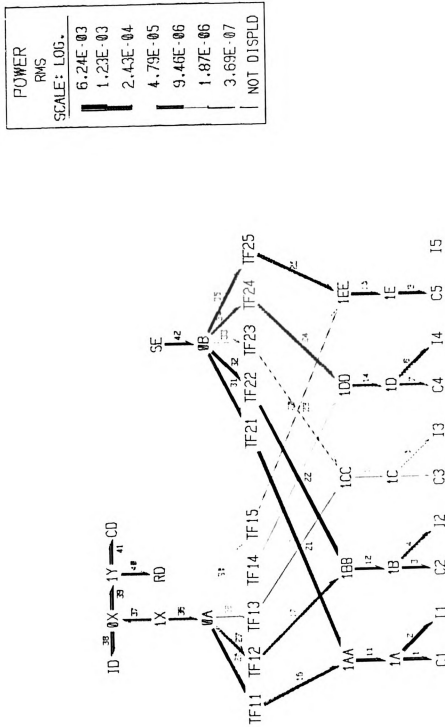


Figure 4-13 Power distribution in the modal bond graph model

the beam and consists of five modal pairs of C and I nodes. Each pair may be treated as a subsystem. The subsystem 1 composed by C1 and I1, and so on. These subsystems connect with the mother system by bonds 11, 12, 13, 14, and 15, respectively. These bonds are cut bonds and consist a cut set. We can apply the connection index CI to judge the significance of each individual subsystem. First, we compute the connection indices as follows

$$\begin{aligned}
 CI_1 &= \bar{W}_{11} / (\bar{W}_{11} + \bar{W}_{12} + \bar{W}_{13} + \bar{W}_{14} + \bar{W}_{15}) \\
 &= 0.005808 / (0.005808 + 0.001931 + 0.000007735 + 0.000591 + 0.0003871) \\
 &= 0.666 \\
 CI_2 &= 0.221 \\
 CI_3 &= 0.000887 \\
 CI_4 &= 0.0677 \\
 CI_5 &= 0.0440
 \end{aligned}$$

By comparison of the obtained connection indices in Table 7, we found that the third mode represented by subsystem 3 has the weakest connection. Therefore, removal of mode 3 should not have much effect on the performance of the whole system. We may try to remove mode 5 and 4, sequentially.

Table 4-6 The powers on cut bonds

ORDER	BOND	RMS POWER	CI %
1	11	0.5808E-02	66.567
2	12	0.1931E-02	22.131
3	14	0.5910E-03	0.887
4	15	0.3871E-03	6.774
5	13	0.7753E-05	4.436

If a model simplification that ignores the highest three modes is used (See Figure 4-14), then the resulting RMS power distribution in list format is given in Table 4-7. The eigenvalues of the original and the reduced models are listed in Table 4-8. To further investigate the behavior of the simplified model as compared to the original model we plotted the velocities at the input (F.42) and the load (F.37) locations versus time in Figure 4-15 ('F' denotes flow, hence velocity in ENPORT). The curves for reduced model are more smooth. The deviations are quite small. This example shows that the weak connection of a subsystem to the mother model is a good indicator to lead to a model simplification.

The computational effort is reduced dramatically by the model reduction. The data in Table 4-9 were obtained by running the models in a Prime 750 system using ENPORT7-7.1.5. The computation time is decreased by reducing the number of state variables and eliminating the high frequency modes. It is most clear when both the fourth and fifth modes were eliminated.

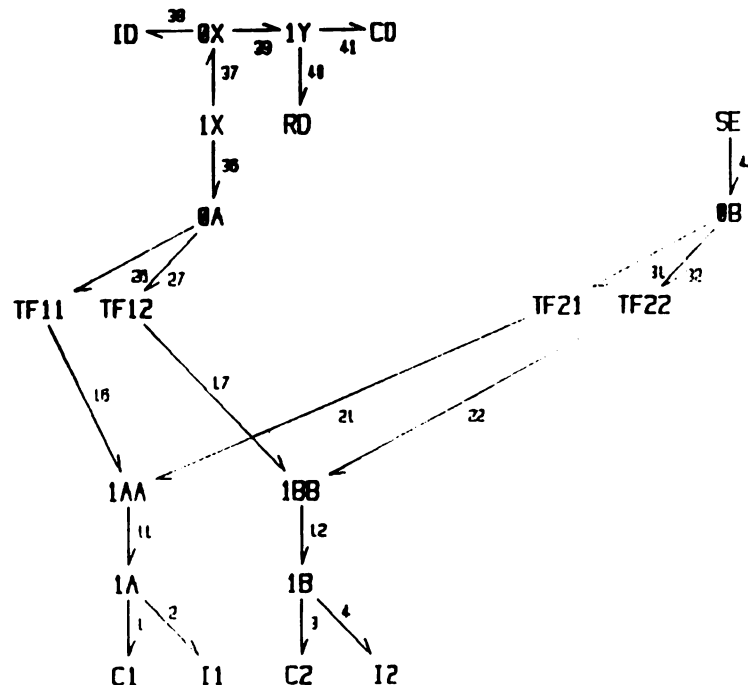


Figure 4-14 Reduced model with two modes

Table 4-7 The power listing of two-mode model

ORDER	BOND	RMS POWER	EI %
1	1	0.6248E-02	100.000
2	42	0.6147E-02	98.379
3	11	0.5817E-02	93.101
4	21	0.5792E-02	92.698
5	31	0.5792E-02	92.698
6	2	0.2154E-02	34.469
7	3	0.2051E-02	32.822
8	12	0.1913E-02	30.626
9	32	0.1875E-02	30.010
10	22	0.1875E-02	30.010
11	38	0.1582E-02	25.316
12	39	0.1486E-02	23.782
13	41	0.1398E-02	22.374
14	36	0.8515E-03	13.628
15	37	0.8515E-03	13.628
16	4	0.7340E-03	11.748
17	16	0.7197E-03	11.518
18	26	0.7197E-03	11.518
19	40	0.4739E-03	7.586
20	17	0.3871E-03	6.196
21	27	0.3871E-03	6.196

Table 4-8 Eigenvalues of the original and reduced models

Original	Reduced
0.00000E+00 ± j 2.46739E+02	
-1.40411E-01 ± j 1.57952E+02	
-3.75707E-01 ± j 8.90238E+01	
-4.12466E-01 ± j 3.99433E+01	-4.17791E-01 ± j 3.99438E+01
-8.06808E-01 ± j 1.17678E+01	-8.17999E-01 ± j 1.17833E+01
-2.64687E-01 ± j 8.24499E+00	-2.64250E-01 ± j 8.25372E+00

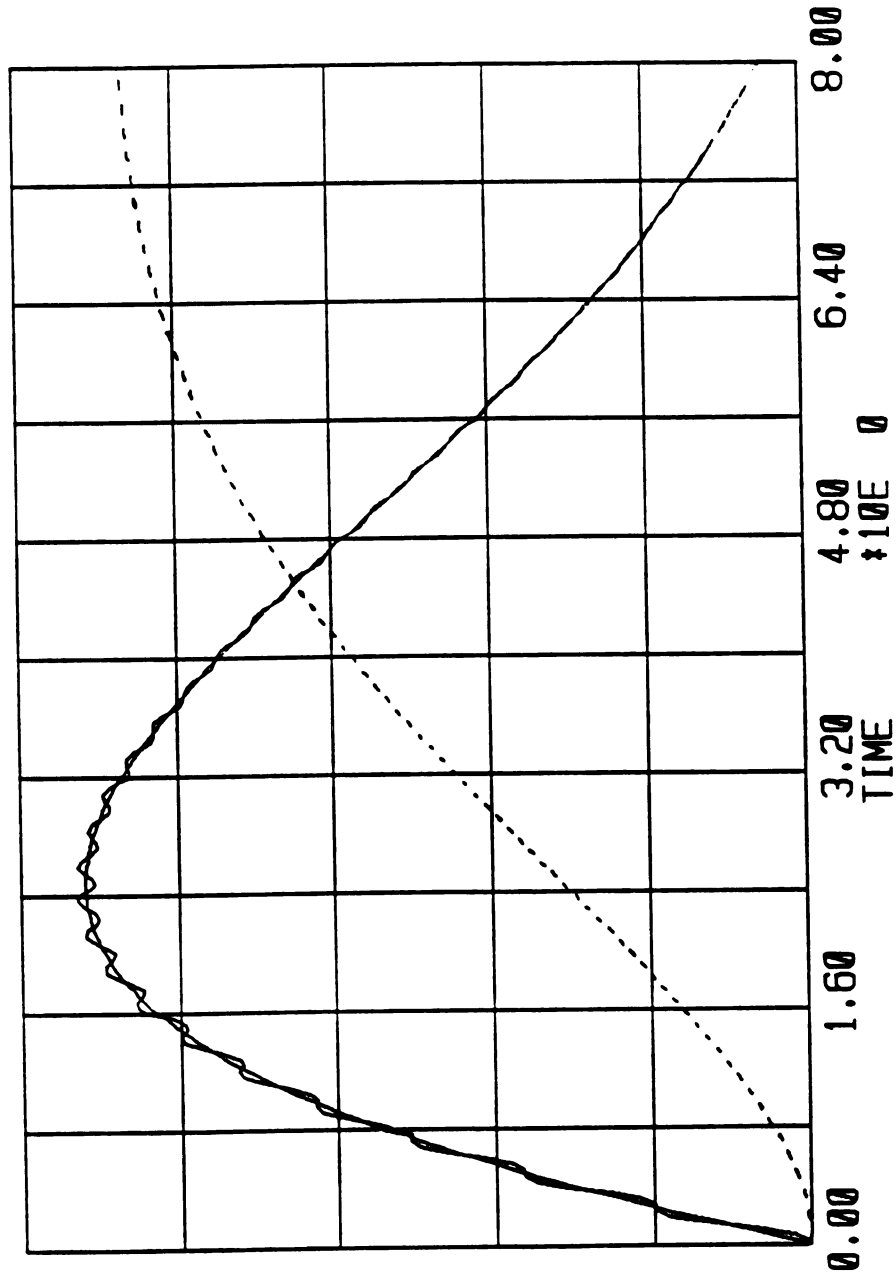
Table 4-9 Computer times

Modes in model	CPU time (sec)
1,2,3,4,5	27.809
1,2,4,5	20.488
1,2,4	13.591
1,2	2.955

SCALING

W
 2.50E-01
 0.00E+00

FB
 1.25E+00
 0.00E+00



LEGEND: W — FB ---

Figure 4-15 Comparison of input and load velocities.

From the examples above the blue color on the thinnest bonds indicates the lowest power level and the user can easily spot those bonds qualitatively, hence the least significant physical effects. Referring to the power listing, the user obtains the power levels quantitatively. Such process is very natural and intuitive. The simplification may be done repeatedly by eliminating the lowest power bond and associated node, or subsystems. In fact, the beam-load model experienced three times of reduction, each time a pair of modal C and I nodes were removed.

Note here that the two examples are linear and time-invariant. It is for the purpose of making a comparison by using the system eigenvalues. It is observed that in this power-based method there is no functional analysis or matrix transformation involved. The method is built on the solution data obtained in simulation, therefore, no restriction exists with respect to the linearity of the system in the problem.

It is clear that since this method must use the simulation data to find the power distribution, we can not escape the initial computation for solving the original high-order, complex, or ill-modeled system model. It may prevent the application in the case that the computer does not have enough storage space for large number of solution data.

Chapter 5

SUMMARY AND CONCLUSIONS

In this dissertation two approaches to the design assessment and simulation of engineering systems were presented. The principal research objective of providing tools for improving insight and efficiency in dynamic system design was achieved.

5.1 Implicit R-field Simulation

With respect to the simulation of models containing implicit R-fields, our research has accomplished the following.

1. For a bond graph containing IRFs, the IRFs were identified at the graph level from causality data. Each IRF was identified separately. The total computation effort was reduced by this localization of IRFs. The identification procedure was implemented in software.

2. A new algorithm that finds the minimum iteration set of a given IRF was presented. This algorithm uses a junction structure property, the basis order, to determine the minimum number of iteration variables and applies to IRFs with a simple junction structure and 1-port R nodes. The minimum number is equal to the minimum of the pair of basis order numbers, E and F. This number is always less than or equal to one half the number of bonds on the IRF R nodes.

3. To extend the above algorithm to an IRF containing multiport R nodes and a simple junction structure, a rule to determine the minimum number of iteration variables for a given causal IRF was deduced from an equivalent digraph. The coupling of port variables by multiport R nodes increases the size of the minimum iteration variable set from $\min(E, F)$ to $\min(E, F) + \sum_{m=1}^M \min(N_r, N_g)_m$. The summation term reflects the couplings within multiport R-nodes.

4. Next an acausal IRF with a multiport node and a simple junction structure was considered. The question was "What assignment of causality will lead to a minimum iteration variable set?". A strategy for assigning causality such that the number of iteration variables for an IRF containing one multiport R node is a minimum was developed. This strategy also determines such a set of iteration variables.

5. The algorithm described above utilizes the property of effort-to-effort and flow-to-flow transformations of simple (and weighted) junction structures. For general junction structures (containing one or more essential gyrator nodes) we developed a new result that gives the basis order. The basis order numbers, E and F, are not a unique pair. In order to find the range of basis order numbers, we transformed a general junction structure to an associated gyrograph. By using the gyrograph representation and the graph matching concept, an algorithm for determining the F-maximum and the F-minimum of basis order for a general junction structure was developed. The structure of the basis set was determined to range from $F_{\min}$ to $F_{\max}$ by steps of two in F.

The author makes the following suggestions for future research on this topic.

1. The restriction on the number of multiport R node to ones in the algorithm for optimum causality assignment in IRFs should be relaxed. The interactions among multiport R nodes must be considered.

2. Apply the basis order algorithm for general junction structure to the solution of IRFs with a general junction structure.

3. Implement the algorithms from 1 and 2 above within a bond-graph-based simulation program, such as ENPORT-7. It will make the software more powerful and efficient in design assessment and simulation.

5.2 Power Distribution and Display

To increase insight and reduce simulation time an approach to model order reduction and simplification based on the power distribution response of dynamic systems was investigated. A new display tool was developed and implemented in software. A number of examples were provided to illustrate the use of the tool. To obtain power distribution attributes the engineering system is modeled by a bond graph, a simulation is made under a set of given test conditions, and time histories of system variables, including powers, are obtained. Various measures of the power response can be calculated and displayed directly on the bond graph by color coding. Such displays give engineers an easy way to visualize power distribution attributes in large-scale, complex dynamic systems under a given set of operating conditions. Often insight may be gained into possibilities for model simplification.

A major advantage of the approach is that it is associated with the graph model and does not require the engineer to deal with equations. The power response approach is not restricted to linear models. In the nonlinear case there are no convenient data of the class of eigenvalues to which to appeal for insight. Here the power distribution may offer an efficient way to make a trial assessment of possible model simplification.

The following issues remain for future development with respect to the power distribution method.

1. It will be useful to get the existing tools for power response assessment, now implemented in pilot version, into wide use so that data regarding modeling approximations can be accumulated. From such shared experience it may be possible to find guidelines that are automatable and will provide the engineer with additional insight.

2. We have observed that different power measures, such as mean or RMS, can yield different insight into the relative importance of parts of the model. With respect to possible model simplification we are at the "cut and try" stage. It would be valuable to have some guidelines as to how to exploit power level distribution data with respect to system response. One such question is "What different measures should be considered for different physical effects, such as power sources, power dissipation, and energy storage?"

3. Power distribution properties may open a new research direction in control. For example, are there some ways to control a dynamic system based on not only signals but also power properties? To answer this

question we need to study the power properties in a dynamic system theoretically. The relation of power variables to many important aspects in control, such as controllability, observability, and stability of a system, need to be investigated. One may start with a simple linear time-invariant model.

APPENDIX

APPENDIX : Standard Bond Graphs and Gyrobondgraphs

1. Standard bond graphs

A standard bond graph consists of elements of the standard set (Table A-1): C, I, R, Se, Sf, 0, 1, TF and GY, which are called capacitance, inertance, resistance, source of effort, source of flow, one-junction, zero-junction, transformer and gyrator, respectively. The elements C, I, R, Se and Sf represent the energy field effects of lumped-parameter multiport physical components, while the elements 0, 1, TF, and GY are used to form the junction structure of the bond graph. The junction structure defines the connection pattern among the field multiports, and it is power-conserving. The SJS contains only elements 0 and 1, while the WJS contains the elements 0, 1, and TF. They share a common causality structure, in the sense that the effort transforms to effort and flow transforms to flow. The efforts and flows are disjoint. Their causal properties are clearly established. The GJS contains gyrator (GY) as well as the elements 0, 1, and TF. The properties of GJS have not been established.

2. Gyrobondgraphs

The gyrator is a fundamental element in constructing power dual pairs of elements, such as (C, I), (Se, Sf) and (0, 1). Based on the gyrators, the standard set can be reduced to a smaller working set (I, R, Se, 1 and GY). Bond graphs can be transformed into a canonical form, gyrobondgraph, from which the properties of the original bond graph can be deduced. For example, see Rosenberg [9]. In this paper, we will use

the set (I, R, Se, l and GY) as our primitive set for expressing the gyrobondgraph.

A standard bond graph can be transformed to its associated gyrobondgraph by the following procedure:

- (1) Replace each standard element not in the primitive set by its equivalent:
 - (a) an Sf is equivalent to an Se and a GY;
 - (b) a C is equivalent to an I and a GY;
 - (c) a TF is equivalent to two GYs in cascade (one of which has unity modulus), and
 - (d) a 0 is equivalent to a l and a GY at each port.
- (2) Eliminate all pairs of unit gyrators.
- (3) Combine all l-junctions that are directly bonded.
- (4) Insert l-junctions to meet the adjacency conditions, namely,
 - (a) each I, R, and Se is adjacent to a l-junction; and
 - (b) each GY is adjacent to two distinct l-junctions.
- (5) Combine any fragments of the graph that require it into equivalent I- or R- fields.

3. Gyrographs

The word, gyrograph, was first used by Professor H. M. Paynter in 1968. He stated that all reduced gyrobondgraphs contain ideal multiport junction structures consisting solely of interconnected l's and GY's, and all such systems are but specializations of a general ideal multiport we call gyrostructure. A graph, simplified from its gyrobondgraph is called the gyrograph.

To simplify a gyrobondgraph to a gyrograph the following steps are used

- (1). Replace each 1-junction by a corresponding circle.
- (2). Replace each gyrator together with its bonds by a edge joining the two adjacent circles.
- (3). Replace each external node by a square.

Table A-1 Standard node sets

<u>SYMBOL</u>	<u>DEFINITION</u>	<u>NAME</u>
SE $\xrightarrow{e}$	$e = e(t)$	source of effort
SF $\xrightarrow{f}$	$f = f(t)$	source of flow
C $\xleftarrow[e]{f}$	$e = \Phi(q)$ $q(t) = q(t_0) + \int f \cdot dt$	capacitance
I $\xleftarrow[e]{f}$	$f = \Phi(p)$ $p(t) = p(t_0) + \int e \cdot dt$	inertance
R $\xleftarrow[e]{f}$	$\Phi(e, f) = 0$	resistance
$\xrightarrow{1} \text{TF} \xrightarrow{2}$ 1:m	$e_1 = m \cdot e_2$ $m \cdot f_1 = f_2$	transformer
$\xrightarrow{1} \text{GY} \xrightarrow{2}$ r	$e_1 = r \cdot f_2$ $e_2 = r \cdot f_1$	gyrator
$\xrightarrow{1} \text{O} \xrightarrow{3}$ $\uparrow 2$	$e_1 = e_2 = e_3$ $f_1 + f_2 - f_3 = 0$	common effort junction
$\xrightarrow{1} \text{I} \xrightarrow{3}$ $\uparrow 2$	$f_1 = f_2 = f_3$ $e_1 + e_2 - e_3 = 0$	common flow junction

BIBLIOGRAPHY

BIBLIOGRAPHY

Aoki, M., "Control of Large-scale Dynamic Systems by Aggregation," IEEE Trans. Automatic Control, AC-13, pp. 240-253, 1968.

Bellman, R., "Vector Lyapunov Functions," SIAM Journal of Control, Vol. 1, pp. 32-34, 1962.

Bos, A. M. and Breedveld, P. C., "1985 Update of the Bond Graph Bibliography," J. Franklin Institute, Vol. 319, No. 1/2, pp. 269-286, 1985.

Breedveld, P. C., "Essential Gytrators and Equivalent Rules for 3-port Junction Structures," Journal of Franklin Institute, Vol. 318, No.2, pp. 77-89, August 1984.

Burreto, J., and Lefevre, J., "R-fields in the Solution of Implicit Equations," J. of the Franklin Inst., Vol. 319, No 1/2, pp. 227-237, 1985.

Chen, C. F., and Shieh, L. S., "Continued Fraction Inversion by Routh's Array," IEEE Trans. Circuit Theory, CT-16, pp. 197-202, 1969.

Davison, E. J., "A Method for Simplifying Linear Dynamic Systems," IEEE Trans. Automatic Control, AC-12, pp. 119-121, 1967.

Dauphin-Tanguy, G., Borne, P. and Lebrun, M., "Order Reduction of Multi-Time Scale Systems Using Bond Graphs, The Reciprocal System and Singular Perturbation Method," Journal of the Franklin Institute, Vol. 319, No.1/2, pp. 157, 1985.

Decarlo, R. A., and Saeks, R., Interconnected Dynamical Systems, Marcel Dekker Inc., New York, 1981.

Van Dixhoorn, J. J., "Simulation of Bond Graphs on Minicomputers," J. Dynam. Syst., Meas. & Control, Vol 99, pp. 9-14, 1977.

Granda, J. J., "Bond Graph Modeling Solutions of Algebraic Loops and Differential Causality in Mechanical and Electrical Systems," Proc. Applied Simulation and Modeling, pp. 188-193, IASTED Conference, San Francisco, CA, 1984.

Granda, J. J., "Computer Generation of Physical System Differential Equations Using Bond Graphs," J. Franklin Inst., Vol 319, No.1/2, pp. 243-255, 1985.

Guardabassi, G., and Locatelli, A., Large Scale Systems: Theory and Applications, IFAC, Pergamon Press, Oxford, 1976.

Happ, H. H., Diakoptics and Networks, Academic Press, New York, 1971.

Harrison, B. K., "Large-scale Linear Systems, in: System Theory," L. A. Zadeh and E. Polak (Eds.), McGraw Hill, New York, 1972.

Hood, S., Rosenberg, R., Withers, D., and Zhou, T., "An algorithm for Automatic Identification of R-fields in Bond Graphs," IBM Journal of Research and Development, Vol. 31, No. 3, pp. 382-390, May 1987.

Hrovat, D., Tobler, W., and Tsangarides, M., "Bond Graph Modeling of Dominant Dynamics of Automotive Power Trains," Dynamic Systems: Modeling and Control, ed., M. Donath, ASME DSC-Vol. 1, 1985.

Jamishidi, M., Large-Scale Systems: Modeling and Control, North-holland, New York, 1983.

Karnopp, D. C. and Rosenberg, R. C., "Analysis and Simulation of Multiport Systems," MIT Press, Cambridge, MA, 1968, See Section 4.3.

Karnopp, D. C., "Power-Conserving Transformations: Physical Interpretations and Applications Using Bond Graphs," J. Franklin Institute, Vol. 288, No. 3, pp. 175-201, 1969.

Karnopp, D. C., "Correspondence on Bond Graphs and 'The Penetrating but Olympian Tensorial Methods of Kron'," The Matrix and Tensor Quarterly, pp. 120-124, June 1970.

Kawaji, S., and Shiotsuki, T., "Model Reduction By Walsh Function Techniques," Mathematics and Computers in Simulation, Vol. 27, North-Holland, 1985.

Kokotovic, P. V., Perkins, W. R., Cruz, J. B., and D'Ans, G., " ϵ -coupling Method for Near-optimum Design of Large-scale Linear System," Proc. IEE 116, pp. 887-892, 1969.

Kokotovic, P. V., Khalil, H., and O'Reilly, J., Singular Perturbation Methods in Control: Analysis and Design, Academic Press Inc., Orlando, Florida, 1986.

Kron, G., Diakoptics, The Piecewise Solution of Large-scale Systems, Macdonald, London, 1963.

Lorenz, F., and Wolper, J., "Assigning Causality in the Case of Algebraic Loops," J. of the Franklin Inst., Vol. 319, No. 1/2, pp. 237-243, 1985.

Margolis, D. L., and Tabrizi, M., "Reduction of Models of Interacting Lumped and Distributed System Using Bond Graphs and Repeated Model Decomposition," J. Franklin Institute, Vol. 317, No. 5, pp. 309-322, 1984.

Matrosov, V. M., "On the Theory of Stability of Motion (in Russian)," Prikladaia Matematika i Mekhanika, 26, pp. 992-1002, 1962.

Meirovitch, L., Analytical Methods in Vibrations, The MacMillan Co., New Yourk, NY, 1967.

Michel, A. N., and Miller, R. K., Qualitative Analysis of Large Scale Systems, Academic Press, New York, 1977.

Ort, J. R., and Martens, H. R., "The properties of Bond Graph Junction Structure Matrices," ASME, Journal of Dynamic Systems, Measurement, and Control, Vol. 95, No. 4, pp. 362-367, December 1973.

Paynter, H. M., "The Rational Foundations of Multiport Dynamics," July 27, 1967 (unpublished notes)

Perelson, A. S., "Bond Graph Junction Structures," ASME, Journal of Dynamic Systems, Measurement, and Control, Vol. 97, No. 2, pp. 189-195, June 1975.

Pernebo, L., and L. M. Silverman, "Model Reduction via Balanced State Space Representations," IEEE Trans. Automatic Control, AC-27, No. 2, pp. 382-387, 1982.

Prakash, R., Palmer, E., and Withers, D., "Design assessment: Electro-Mechanical Systems," Research Report RC-10993, IBM Thomas J. Watson Research Center, Yorktown Heights, NY, 1985.

Rosenberg, R. C., "State-Space Formulation for Bond Graph Models of Multiport Systems," ASME Journal of Dynamic Systems, Measurement, and Control, Vol. 93 NO. 1, pp. 35-40, March 1971.

Rosenberg, R. C. and Karnopp, D., "A Definition of the Bond Graph Language," ASME, Journal of Dynamic Systems, Measurement, and Control, Vol. 94, No. 3, pp. 179-182, Sept. 1972.

Rosenberg, R. C., "On Gyrobondgraphs and Their Uses," ASME, Journal of Dynamic Systems, Measurement, and Control, Vol. 100, No. 1, pp. 76-82, March 1978.

Rosenberg, R. C., "Essential Gytrators and Reciprocity in Junction Structure," Journal of Franklin Institute, Vol. 308, No. 3, pp. 343-352, September 1979.

Rosenberg, R. C., and A. N. Andry, Jr., "Solvability of bond Graph Junction Structures With Loops," IEEE, Transaction on Circuits and Systems, Vol. CAS-26, No.2, pp. 130-137, February 1979.

Rosenberg, R. C., and Moultrie, B., "Basis Order for Bond Graph Junction Structures," IEEE Transaction on Circuits and Systems, CAS-27, No. 10, pp. 909-920, Oct. 1980.

Rosenberg, R. C., and Karnnop, D. C., Introduction to Physical System Dynamics, McGraw-Hill, New York, 1983.

Rosenberg, R. C., The ENPORT-5 User's Manual, ROSENCODE Associates, Lansing, MI, 1984.

Rosenberg, R. C., "Improving Design Assessment Techniques for Reliability and Yield by Bond Graph Technology", Proposed to IBM Manufacturing Research, Department of Mechanical Engineering, Michigan State University, 1985.

Rosenberg, R. C., The ENPORT-7 Reference Manual, ROSENCODE Associates, Lansing, MI, 1986.

Sage, A. P., Methodology for Large Scale Systems, McGraw-Hill, New York, 1978.

Sandell Jr., N. R., Varaiya, P., Athans, M., and Safonov, M. G., "Survey of Decentralized Control Methods for Large Scale Systems," IEEE Trans. on Automatic Control, Vol. AC-23, No. 2, April 1978.

Siljak, D. D., Large-scale Dynamic Systems, North-Holland, New York, 1978.

Siljak, D. D., "Complex Dynamic Systems: Dimensionality, Structure, and Uncertainty," Large Scale Systems 4 (1983), pp. 279-294.

Skelton, R. E., "Cost Decomposition of Linear Systems with Application to Model Reduction," Int. J. Control, Vol. 32, No. 6, pp. 1031-1055, 1980.

Skelton, R. E., and Yousuff, A., "Component Cost Analysis of Large Scale Systems," Int. J. Control, Vol. 37, No.2, pp. 285-304, 1983.

Syslo, M. M., Deo, N., and Kowalik, J. S., Discrete Optimization Algorithms, Prantice Hall, Englewood Cliffs, NJ, 1983, pp. 320.

Titli, A., and Singh, M. G., Large Scale Systems: Theory and Applications, IFAC, Pergamon Press, Oxford, 1980.

Thompson, W. T., Theory of Vibrations with Applications, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1981.

Tse, E. C. Y., Medanic, J. V., and Perkins, W. R., "Chained Aggregation of Linear Time Invariant Systems," Proc. JACC. San Fransisco, CA. 1977.

Yousuff, A., and Wagie, D. A., and Skelton, R. E., "Linear System Approximation via Covariance Equivalent Realization," J. Math. Analysis and Appl., 106, pp. 91-115, 1985.

Zalewski, Z., and Rosenberg, R. C., "Simulation of Engineering Models Containing Bond Graphs and Block Diagrams," ASME Computer in Engineering Conference, 355-359, Chicago, IL, July 1986.

Zeid, D. A., "Bounds on the Eigenvalues for Certain Classes of Dynamic Systems," Ph. D Dissertation, Michigan State University, 1982.

Zeid, D. A., "Solution of a set of Nonlinear Algebraic Equations for General-Purpose CAD Program," IEEE Circuits and Devices Magazine, pp. 7-20, Sep. 1985.

Zhou, T., "A Parallel Computation Method for Dynamic Systems with Coupled nonlinear Dissipation," MS Thesis, Michigan State University, 1984.