

Structural Analysis from System Configurations for Modeling and Design of Multi-Energy Domain Dynamic Systems

by

Shih-Ying Huang

B.S., Mechanical Engineering (1988)

National Taiwan University

S.M., Mechanical Engineering (1993)

Massachusetts Institute of Technology

Submitted to the Department of Mechanical Engineering
in Partial Fulfillment of the Requirements for the degree of
Doctor of Philosophy in Mechanical Engineering

at the

Massachusetts Institute of Technology

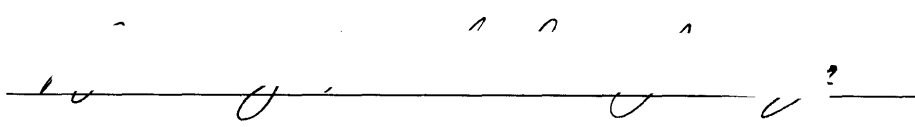
May, 1997

©Shih-Ying Huang 1997

All rights reserved

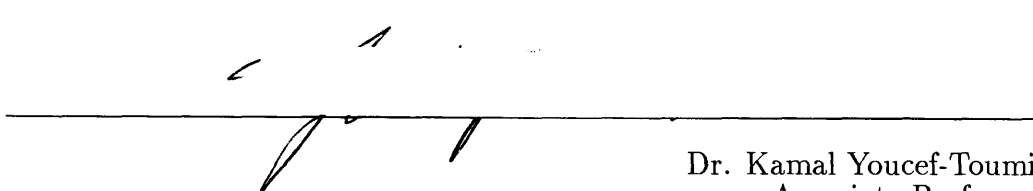
The author hereby grants to MIT permission to reproduce and to distribute publicly copies of this thesis document in whole or in part.

Signature of Author




Department of Mechanical Engineering
May 23, 1997

Certified by



Dr. Kamal Youcef-Toumi
Associate Professor
Thesis Supervisor

Accepted by



MASSACHUSETTS INSTITUTE
OF TECHNOLOGY

Dr. Ain A. Sonin
Chairman, Department Committee on Graduate Students

JUL 21 1997

Eng.

Structural Analysis from System Configurations for Modeling and Design of Multi-Energy Domain Dynamic Systems

by

Shih-Ying Huang

Submitted to the Department of Mechanical Engineering
on May 23, 1997 in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in
Mechanical Engineering

ABSTRACT

Modeling and design of dynamic systems play a major role in determining the closed-loop system performance. However, many design methodologies still rely on trial-and-error procedures and numerical simulations. Due to the lack of physical insights, these approaches not only defer the conceiving of better system configurations but also may lead to unnecessary loss of efficiency. On the other hand, the study of dynamic interactions provides a valuable guidance for achieving proper system behaviors. In this research, an energy-based method is used to provide a unified representation for multi-energy domain systems. By the use of its energy interactions and causality implications, it is possible to determine the inherent system properties early in the design stage before detailed element characteristics and equations are determined. The obtained information in turn suggests feasible directions for design improvement toward better system performance. Three main topics are presented in this thesis to demonstrate the capability of the proposed structural analysis procedures.

The first topic addresses the problem of excess states and their influences to the system analysis procedures. The excess states usually exist in certain over-constrained structures. It is found that by using explicit field representations, such ambiguities can be eliminated. Based on this approach, a set of model revision procedures are developed to eliminate the excess states so that the existing analysis procedures can be properly applied.

The second topic is the identification of relative degrees and zero dynamics for non-linear MIMO systems. A method is proposed to derive the zero dynamics of physical systems from bond graph models. This method incorporates the definition of zero dynamics in the differential geometric approach and the causality manipulation in the bond graph representation. By doing so, the state equations of the zero dynamics can be easily obtained. The system structure and elements which are responsible for the zero dynamics can be identified. In addition, if isolated subsystems which contribute to the zero dynamics exist, they can be found. Thus, the design of physical systems including the consideration of the zero dynamics becomes straightforward.

The purpose of the third topic is to build the direct relations between the component characteristics and the system eigenvalues. In this thesis, several decomposition procedures are proposed to identify the physical components which contribute most to a certain group of eigenvalues. By using the available matrix theories, the bounds of each eigenvalue group can be represented in terms of the component characteristics. These bounds will facilitate the design of physical systems.

From the results of this research, it is shown that the analysis and design of dynamic systems can be conducted in a systematic way by studying the system configurations. The proposed procedures can be easily coded and become part of a computer-aided design package.

Thesis Committee:
Professor Neville Hogan
Professor Stephen, D. Senturia
Professor Kamal Youcef-Toumi, Chairman

Acknowledgement

First, I would like to thank my thesis supervisor, Professor Kamal Youcef-Toumi for his advice and assistance throughout my thesis work. I also want to thank Professor Hogan and Professor Senturia for serving on my thesis committee. They have devoted considerable time and attention to my thesis work and offered helpful suggestions and comments.

I thank the secretaries and other MIT employees who have ever helped me, especially our department secretary Lesli. Her kindness is well-known among the ME students.

I am thankful to Toyota company for supporting me through research assistantships. I thank our visiting engineer from Toyota, Mr. Senga for his friendship and his help to the project.

I thank the students I have worked with in the flexible automation laboratory. The former and the current students, Tarzen, Jose, T.J., Mitchell, Francis, C.-J., Tetsuo, Doug, and Jake have cheered up my graduate student life. I would like to particularly thank Tetsuo and his wife Renee. Their encouragement has been a great support to me.

I would like to thank my friends, Jory Tsai and his wife, Tung-Ching, Yu-Huai, and Li-Hsiang for the joy and the help that they have brought me. I thank my brother, Yung-Jung, for his constant concern from overseas. I also would like to thank my sister, Chun-Chen. She has been the best companion to me over these years.

Finally, I would like to thank my parents. It is their love and supports that carried me through the difficult times. This thesis is dedicated to them.

Contents

1	Introduction	1
1.1	Motivation	1
1.2	Background	3
1.2.1	Structural Analysis	3
1.2.2	Models Using Energy Methods - Bond Graph Representation .	3
1.2.3	Software Packages	5
1.3	Scope and Contents of the Thesis	6
2	Explicit Fields and Their Application to Structural Property Inspection of Physical Systems	9
2.1	Introduction	9
2.2	Example	11
2.3	Properties of Explicit Energy Storage Fields	13
2.4	Model Revision for Structural Property Inspection	17
2.4.1	The existence of type 1 excess states	17
2.4.2	The use of explicit fields	18
2.4.3	The existence of type 2 excess states	20
2.5	Applications	20
2.5.1	Independent state variables	20
2.5.2	Structural property examination	21
2.5.3	The equilibrium states	22
2.6	Computer Implementation	23

2.7	Conclusion	26
3	Zero Dynamics of Physical Systems from Bond Graph Models - Part I : SISO Systems	27
3.1	Introduction	27
3.2	Zero Dynamics in the Differential Geometric Approach	29
3.3	Zero Dynamics of SISO Systems from Bond Graph Models	31
3.3.1	Definitions	32
3.3.2	Relative degree	33
3.3.3	Zero dynamics	34
3.4	Applications	42
3.4.1	Systems with the input and the output on the same bond	42
3.4.2	Systems with simple structures	43
3.4.3	A Design Example : Prosthesis Arm Design	45
3.5	Conclusion	49
4	Zero Dynamics of Physical Systems from Bond Graph Models - Part II : MIMO Systems	50
4.1	Introduction	50
4.2	Zero Dynamics in the Differential Geometric Approach	51
4.3	Zero Dynamics of MIMO Systems from Bond Graph Models	55
4.3.1	Vector relative degree	55
4.3.2	Dynamic extension procedure	56
4.4	MIMO Systems with a Vector Relative Degree	58
4.5	MIMO Systems without a Vector Relative Degree	62
4.5.1	Systems with a relative degree under dynamic extension	62
4.5.2	Systems with ill-designed input-output configurations	69

4.6	Conclusion	77
5	Decomposition of Linear Dynamics in the Physical Domain and Eigenvalue Estimations	78
5.1	Introduction	78
5.2	Currently Available Methods	79
5.3	Decomposition of Fast-Slow Dynamics	83
5.3.1	Singular perturbation theory	84
5.3.2	Decomposition in the physical domain	84
5.3.3	A numerical example	87
5.4	Decomposition of High-Low Frequency Oscillation Modes	88
5.4.1	An auxiliary transformation	89
5.4.2	Physical interpretations	90
5.5	Decomposition of Heavily-damped and Lightly-damped Dynamics	94
5.5.1	The decomposition procedure	95
5.5.2	A numerical example	100
5.6	Eigenvalue Estimation for General Systems	102
5.6.1	Undecomposable systems	103
5.6.2	A pre-conditioning procedure	105
5.7	Design Examples	107
5.7.1	A mechanical structure	108
5.7.2	An arm prosthesis design	111
5.8	Conclusion	117
6	Conclusion and Recommendations	119
A	Proofs concerning the explicit fields	122

A.1	Independent state variables contributed by explicit fields	122
A.2	The coupling in explicit fields	124
B	Proofs concerning the relative degrees and the zero dynamics	129
B.1	Relative degrees	129
B.2	Supplements to the ZDIP procedure	132
B.3	The vector relative degrees	132
C	A maple procedure	136

List of Figures

2.1	Schematic of an inductive sensor.	12
2.2	Bond graph models of an inductive sensor.	12
2.3	Revised bond graph models.	13
2.4	A C field with mixed causality assignment.	14
2.5	Bond graph models of a magnetic bearing.	21
2.6	Bond graph model with derivative causality assignment and the revised model.	22
2.7	Equilibrium states from bond graph model with field representation.	23
2.8	A simple example of grouping capacitances.	24
2.9	Bond graph model with a field representation.	25
3.1	A mass-damper-spring unit.	34
3.2	A bond graph model of the mass-damper-spring unit.	35
3.3	A choice of causal path.	35
3.4	The shortest causal path.	36
3.5	The zero dynamics model.	38
3.6	The partitioned zero dynamics model.	41
3.7	(a) Bond graph model of a simple RC system. (b) The zero dynamics.	43
3.8	Bond graph model of a speaker.	44
3.9	The causality assignment after ZDIP procedure.	45
3.10	The zero dynamics identified by ZDIP procedure.	45
3.11	The schematic of an arm prosthesis.	46
3.12	The bond graph model of an arm prosthesis.	46

3.13	The zero dynamics model of an arm prosthesis.	47
3.14	Design 1 of an arm prosthesis.	48
3.15	Design 2 of an arm prosthesis.	48
4.1	Bond graph model of an MIMO system.	60
4.2	The causality assignment after ZDIP procedure.	60
4.3	The zero dynamics identified by ZDIP procedure.	61
4.4	A 2–input 2–output mechanical system.	62
4.5	A corresponding bond graph model.	63
4.6	The shortest causal paths of an MIMO model.	70
4.7	The shortest causal paths of an MIMO model by dynamic extension.	70
4.8	The zero dynamics identified by ZDIP procedure.	71
4.9	A 2-input 2-output mechanical system.	71
4.10	A corresponding bond graph model.	72
4.11	The shortest causal paths of a MIMO model.	76
4.12	The shortest causal paths of a MIMO model by dynamic extension.	76
4.13	An alternative input-output configuration design.	77
5.1	An R-C circuit.	81
5.2	An estimation of eigenvalues from Gersgorin’s theorem.	82
5.3	The first oscillation mode of a mass-spring system.	82
5.4	The highest frequency oscillation mode of a mass-spring system.	83
5.5	The bond graph model of an R-C system.	85
5.6	The bond graph model of the fast dynamics.	86
5.7	The bond graph model of the slow dynamics.	86
5.8	An isolated R-C loop.	86
5.9	A short-circuited I element.	87

5.10	The bounds of the eigenvalues.	88
5.11	The effects of the auxiliary transformation.	91
5.12	An $I - C$ system.	92
5.13	Case 1: high frequency oscillation mode.	92
5.14	Case 1: low frequency oscillation modes.	92
5.15	Case 1: low frequency oscillation modes.	93
5.16	Case 2: high frequency oscillation modes.	93
5.17	Case 2: low frequency oscillation modes.	93
5.18	Case 2: low frequency oscillation modes.	94
5.19	The eigenvalue distribution of the systems with both light and heavy dissipations.	95
5.20	A simple mass-damper-spring system.	98
5.21	The corresponding bond graph model.	99
5.22	The bond graph model representing the heavily-damped modes. . . .	99
5.23	The bond graph model representing the lightly-damped modes. . . .	100
5.24	The equivalent bond graph model representing the lightly-damped modes.	100
5.25	The bond graph model representing the lightly-damped modes. . . .	100
5.26	The decomposition results (a) $\zeta_2 = 0.7$. (b) $\zeta_2 = 1.4$	101
5.27	A summary of the proposed decomposition procedures.	102
5.28	A decomposable distribution of eigenvalues.	103
5.29	An undecomposable distribution of eigenvalues.	104
5.30	The eigenvalue distribution of a Butterworth type filter.	104
5.31	A bond graph model.	105
5.32	A simple partition.	106
5.33	A mechanical structure.	108

5.34	The corresponding bond graph model.	108
5.35	The decomposition of heavily-damped (a) and lightly-damped (b) modes. 109	
5.36	The decomposition of high (a) -low (b) frequency oscillation modes.	109
5.37	The decomposition of fast (a) -slow (b) dynamics.	109
5.38	The estimated eigenvalues from the decomposed subsystems.	111
5.39	The physical systems representing the dominant dynamics.	111
5.40	The oscillation modes for the dominant dynamics.	112
5.41	The bond graph model of an arm prosthesis.	112
5.42	The eigenvalue distribution of the arm prosthesis.	113
5.43	The local loop of C_b, I_f	113
5.44	The local loop of C_b, I_m	114
5.45	The bond graph model representing the pure oscillation eigenmode. .	114
5.46	The equivalent bond graph model representing the pure oscillation eigenmode.	115
5.47	The bond graph model corresponding to the real eigenvalue.	115
5.48	A simplified bond graph model.	116
5.49	A simplified bond graph model.	116
5.50	A simplified bond graph model.	116
5.51	The bond graph model representing the real eigenmode.	117
A.1	A new representation of explicit field with no constraints inside. . . .	127
A.2	(a) The original system. (b) An implicit field form.	128
A.3	(c) The system after releasing the constraints. (d) The final form. . .	128
B.1	A causal path which contains a derivative causality.	130
B.2	An alternative causality assignment.	130

B.3	An alternative causality assignment.	131
C.1	The bond graph model of an arm prosthesis.	136
C.2	The results from the MAPLE procedure.	137

Chapter 1

Introduction

1.1 Motivation

Modelling is a critical step in dynamic system design and feedback system synthesis. Although there are many other important factors in the design processes, the overall performance limitation of the resultant system is usually determined by the use of dynamic models. Sophisticated models surely predict the system behavior well. However, a very detailed and accurate model might not provide much useful insight and information for the design purpose. Efficient models will require the use of simple elements to capture the essential dynamic characteristics with reasonable valid operating regions. In most of the cases, this depends on whether the model structures reflect the essential dynamic interactions. Thus, the study and implementation of the structural analysis of dynamic interactions can provide efficient modelling and design tools. To fulfill this purpose, the following issues must be considered.

First of all, different model representations could give different levels of insights to the system structures. For example, a transfer function describes the input-output relation of a dynamic system, but the internal interactions are totally missing; State equations preserve the details of state interactions, but the connections of physical elements are not obvious from the representation. To closely relate the physical elements with the system dynamic features, the model must preserve the structures of physical connections. Especially for multi-energy domain systems, the transduction between different energy domains should be clearly represented.

Secondly, the models must provide ways of specifying the causal relations. In many theoretical works which address the structures of dynamic systems, the considered systems are described by state equations in a general form. This indeed enlarges the scope of applications. But it also causes the difficulty of relating the theoretical results with the physical elements and parameters. Since the causal relations imply the dependency of state variables and serve as a guidance for equation derivation, they can be adopted to overcome this difficulty if appropriately used.

In terms of implementation, the purpose of the structural analysis is to facilitate the use of existing numerical software packages rather than replace them. Currently, many software packages offer powerful simulation functions and parameter searching tools. However, these functions do not directly inspire useful model structures. The modelling and design processes which utilize these functions still rely on trial-and-error procedures or parameter searching algorithms based on fixed structures. Due to the lack of physical insights, this approach not only defers the conceiving of better system configurations but also possibly leads to unnecessary loss of efficiency. Thus, a pre-analysis of the dynamic interactions before using these functions is necessary.

For the above reasons, one of the energy-based methods - bond graph representation will be used in this thesis to describe the physical connections of multi-energy domain systems. One main goal of this research is to determine the inherent system properties at early design stages before detailed element characteristics and equations are determined. The results can then propose feasible directions for design improvement or suggest useful analysis procedures. By doing so, a better system performance can be achieved by properly designed system structures. Another emphasis of this work is to generate systematic algorithms to automate the structural and other dynamic analysis procedures.

1.2 Background

1.2.1 Structural Analysis

Many important features of dynamic systems are inherently determined by the structure of dynamic interactions. Thus the study of system structures has been an important topic. Theories have been developed for the analysis of dynamic system structures. The results provide systematic calculation procedures for finding system properties. For example, the controllability or observability can be found by checking the rank of controllability/observability matrices. The zero dynamics of general nonlinear systems can be found by a differential geometric approach. However, these approaches provide limited help for system design or model revision. The calculations do not suggest possible solutions for the construction of desired structures. For this reason, certain graphical representations have been used to represent the state interactions and help the calculations. However, it is still difficult to relate the physical elements with the analysis results. As will be explained in the following sections, the model representations based on energy methods are suitable candidates for this purpose.

1.2.2 Models Using Energy Methods - Bond Graph Representation

Bond graph models describe the dynamic behavior of physical systems by connecting idealized lumped elements based on the principle of energy conservation. These network-like models provide very useful insights into the structure of dynamic systems. Causal relations between subsystems can be assigned according to the element characteristics and the junction constraints. Once the bond graph model is built, a set of state variables is easily determined and a set of state equations can be generated systematically [36, 24]. The state variables in bond graph models are directly related

to the energy storage in the system and are easily interpreted from a physical point of view [23].

In the literature, it has been demonstrated that useful information can be extracted from the bond graph representation of physical systems. Important topics which are related to the design and simulation of physical dynamic systems are listed as follows.

Order of systems: [2, 23, 24, 37] pointed out the classes of constraints that influence the order of the systems. [24] introduced a derivative causality assignment procedure to detect a different class of constraints which can not be detected by the integral causality assignment procedure (SCAP procedure). [37] shows a systematic procedure to identify the excess states and the source of constraints (topologically and source induced).

Implicit equations: [3] pointed out that the implicit equations caused by R elements can be identified by the existence of free choice of causality assignment to the R elements. in [46], a general way of identifying the existence of implicit equations is described and defined as zero order path. [50] describes a way of using Lagrange multiplier to eliminate the derivative causality and obtain the implicit equations. [42] uses parasite elements to eliminate derivative causalities and therefore certain class of implicit equations.

Eigenvalues: The problem of obtaining the information about the eigenvalues directly from a bond graph model was also considered in [51, 52, 53]. For a class of systems with uniform parameters, certain bounds of eigenvalues were obtained. The results can be useful for interactive computer-aided system design.

Controllability/Observability: Methods of examining structural properties such

as structural controllability and observability are developed in [41, 43, 44]. Some design methods have been addressed, e.g. how to determine the minimum number of actuators and sensors and their appropriate locations in a physical system [44].

Relative degrees: An inspection rule for the identification of relative degrees in SISO systems was proposed in [47]. For the application of feedback decoupling problems, the vector relative degrees of MIMO systems were studied in [29, 30].

Zero dynamics: Some heuristic rules are found to identify the zeros for a certain class of linear SISO mechanical systems [31]. This can be done directly from the mass-spring-damper schematic. The idea is extended in [47] by using the bond graph representation. For a class of nonlinear SISO systems, the zero dynamics can be obtained by recognizing the junction structure patterns.

Synthesis: Synthesis of a class of linear physical systems which exhibit desired system response by the use of bond graphs is discussed in [34, 35].

1.2.3 Software Packages

Software packages have been developed for the purpose of simulating dynamic systems. These packages may be divided in two categories. The first category deals with simulating mathematical models of systems and possibly providing tools for the analysis and processing of results. Some of these packages include *THTSIM*, *TUTSIM*, *SIMNON*, *MATLAB*, *CONTROL – C*, *MATRIX_x*, *UNYSIS* and *HYCAD*. Such packages provide results and allow the processing of results in the time domain, frequency domain or both domains. The input information can be in the form of line code (transfer functions and state space representation) and/or graphic form (block diagrams). Most of these programs are for general applications and provide a good

environment for simulation and analysis.

Another category deals with packages designed for modelling purposes. These include *ENPORT* and *CAMAS* whose input information is graphical in the form of a *bond graph*. In this case, the software generates the system's equations and simulates the system model. The software has improved in the last few years. Other programs, such as *GEM*, *CAMS* and *MS – BOND*, provide only the state equations of a system. Several other programs based on the bond graph theory have been developed [6]. The characteristics and the applications of most of them are summarized in the survey paper [12]. Some schemes dealing with the generation of symbolic equations are discussed in references [32, 33, 49].

Currently, there is no known software package available which provide diagnostic functions or deal with design (synthesis) and analysis of dynamic systems. Thus, this research will concentrate on the development and implementation of rules, procedures, and algorithms for extracting physical system properties from their graphical models.

1.3 Scope and Contents of the Thesis

Three main topics are presented in this thesis to demonstrate the capability of the proposed structural analysis procedures.

The first topic addresses the problem of excess states and their influences to the system analysis procedures. The excess states usually exist in certain over-constrained structures. In these models, although the representations are legitimate in terms of physical meaning, the resultant excess states cause pitfalls in the inspection of system properties. It is found that by using the explicit field representations, such ambiguities can be eliminated. Based on this approach, a set of model revision procedures are developed to eliminate the excess states so that the existing and the being-developed analysis procedures can be properly applied.

The second topic is the identification of relative degrees and zero dynamics for general nonlinear MIMO systems. Relative degrees and zero dynamics are important features for the design of feedback control laws. For certain systems, the zero dynamics even directly determines the performance limits. Since the intrinsic zero dynamics can not be influenced by any feedback compensation, it is important to design the physical systems so that they possess desired zero dynamics. However, the calculation of the zero dynamics is usually complicated, especially if a form which is closely related to the physical system and suitable for design is required. A method is proposed to derive the zero dynamics of physical systems from bond graph models. This method incorporates the definition of zero dynamics in the differential geometric approach and the causality manipulation in the bond graph representation. By doing so, the state equations of the zero dynamics can be easily obtained. The system elements which are responsible for the zero dynamics can be identified. In addition, if isolated subsystems which exhibit the zero dynamics exist, they can be found. Thus, the design of physical systems including the consideration of the zero dynamics become straightforward.

The purpose of the third topic is to build the direct relations between the component characteristics and the system eigenvalues. It is known that the symbolic solutions for the eigenvalues of high order systems are not available. Even if the exact solutions exist, they may be too complicated, and therefore do not point out useful design directions. In this thesis, several decomposition procedures are proposed to identify the physical components which contribute most to certain group of eigenvalues. By using the available matrix theories, the bounds of each eigenvalue group can be represented in terms of the component characteristics. These bounds will then facilitate the design of physical systems so that they have the eigenvalues roughly at the desired locations.

The thesis is organized as follows: The study of excess states and the model revision procedures are presented in Chapter 2. The identification of zero dynamics for SISO systems is discussed In Chapter 3. The extension of the proposed procedures for MIMO systems are given in Chapter 4. The system decomposition issues and the bounds of eigenvalues are studied in Chapter 5. Finally, concluding remarks are given in Chapter 6.

Explicit Fields and Their Application to Structural Property Inspection of Physical Systems

2.1 Introduction

Bond graph models describe the dynamic behavior of physical systems by connecting ideal lumped elements based on the principle of energy conservation. These network-like models provide very useful insights into the structure of dynamic systems. One major advantage of these graphical notations is the clear representation of constraints, power flow and independent state variables. However, certain over-constrained structures are not dealt with adequately with this approach. Such a situation arises when several energy storage elements of the same type are directly coupled by a junction structure. In these models, although the representations are legitimate in terms of physical meaning, the resultant excess states cause pitfalls in the inspection of system properties. In this chapter, a method using the explicit field representations is proposed to eliminate these ambiguities.

In bond graph representations, elements of the same type can be grouped as a field. According to the definition in [22], the representations of fields with explicit constitutive equations which describe the input-output relation between the field ports are referred to as explicit fields. On the other hand, the representations containing unsolved junction structures are referred to as implicit fields. While both representations are legitimate, the choice of representation would depend on the purpose of

modeling. In an extreme case, all the elements of the same type in the system are grouped as fields. This representation is suitable for a systematic way of deriving state equations. On the contrary, when one examines the structure of a system, the junction connections provide very useful information. Thus the implicit form is preferred. Furthermore, if explicit fields exist in the model, one can use decomposition schemes to represent fields in an implicit form when it is necessary [7]. In this chapter, we will group only certain related elements into an explicit field for the purpose of eliminating the structural ambiguity.

The use of explicit fields has been proposed to eliminate derivative causalities [22, 25] and help the analysis of systems which contains implicit equations [3]. Once the constrained elements are identified and grouped into an explicit field, the revised models usually appear to be clear and compact. In these applications, the grouping scheme is a key issue. Similarly, in order to eliminate the excess states, the identification of the constrained elements need to be addressed first.

By the Sequential Causality Assignment Procedure (SCAP) [22], a set of differential equations in the familiar form $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u})$ can be derived. The constraints between the states \mathbf{x} are represented by the derivative causality in the model. Yet, the constraints between the derivative of the states $\dot{\mathbf{x}}$ remain unidentified. Therefore, this class of constraints influences the results of many structural inspection rules. By an alternative causality assignment procedure [24], from the same model, an alternative set of integral equations in the form $\int^t \mathbf{z} dt = \hat{\mathbf{f}}(\mathbf{z}, \mathbf{u})$ can be derived. The constraints between \mathbf{z} , which corresponds to the variables $\dot{\mathbf{x}}$ in the previous equation, are represented by the integral causality remaining in the model. Thus the possible excess states can be identified. However, due to the imposed sources, this procedure also drops independent states in some cases [23, 24]. In the attempt to avoid such an overestimate of excess states, versatile procedures were developed [2, 37]. Never-

theless, due to the subtlety of this issue, exceptions to these procedures still exist. Also, a systematic way of revising the model for the purpose of structural inspection is still left open. In this chapter, we will refine the procedures of detecting the constraints between \mathbf{z} (or $\dot{\mathbf{x}}$) and derive a systematic model revision procedure by the use of explicit fields. For the convenience of description, in the following contents, the states which are coupled by the constraints in the form of $f(\dot{\mathbf{x}}) = 0$ are defined as type 1 excess states. the states which are coupled by the constraints in the form of $f(\dot{\mathbf{x}}, \mathbf{u}) = 0$ are defined as type 2 excess states.

An example is presented in section 2.2 to point out the issues of excess states. The properties of explicit field representations and their interactions with other elements are discussed in section 2.3. Section 2.4 describes the procedures of utilizing the explicit fields to eliminate excess states. Several examples are presented in section 2.5 to illustrate the use of these procedures. The conclusion is in section 2.6.

2.2 Example

Consider the inductive sensor in Figure 2.1, the input source supplies a reference signal so that the position can be obtained from the measurements V_1 and V_2 . The gaps between the ferromagnetic teeth are modeled as nonlinear capacitances. The corresponding bond graph models are shown in Figure 2.2 (a) when the input is a voltage source and Figure 2.2 (b) when the input is a current source. The causality is assigned using SCAP procedure.

It is well-known that usually the number of integral causalities indicates the number of independent states. Therefore, the number of independent states appears to be 4 in Figure 2.2 (a) and 3 in Figure 2.2 (b) respectively. However, the actual number of independent states in these models are 3 and 2 respectively. This is due to the excess states (type 1) caused by the junction constraints. For simple systems, these

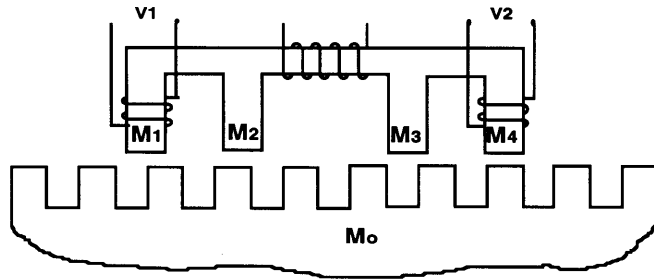


Figure 2.1: Schematic of an inductive sensor.

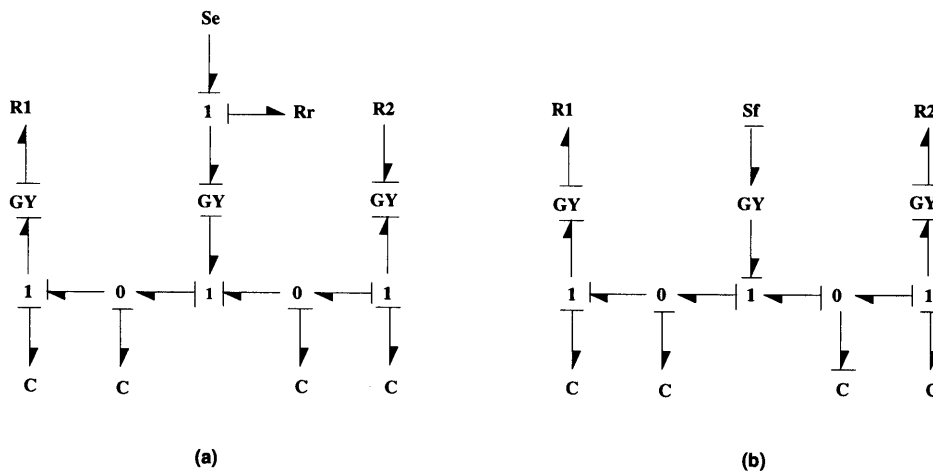


Figure 2.2: Bond graph models of an inductive sensor.

excess states can be easily detected by either deriving the state equations directly or examining the physical system structure. But in general, these excess states can be very confusing, especially if one is interested in the inspection of other structural properties. Since many rules for structural inspection depends on the manipulation of causality and the count of causality, a systematic treatment of this problem is needed. For this example, one can see that if all the capacitances are grouped into a field as shown in Figure 2.3, the number of independent states would be exactly the number of ports which exhibit integral causality. Thus the new model structure is more appropriate for system analysis. This is an example of using the explicit field

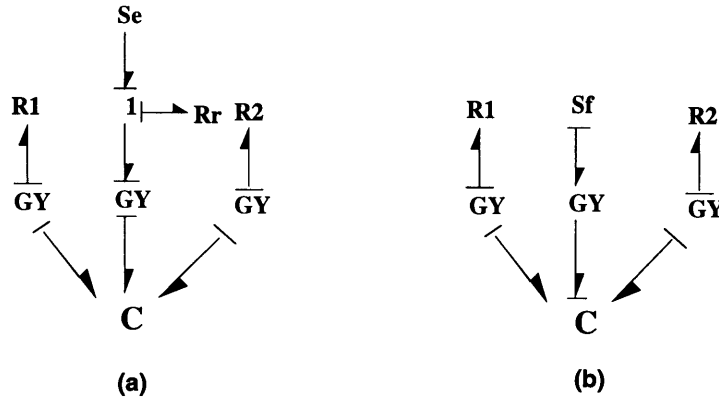


Figure 2.3: Revised bond graph models.

to eliminate excess states. In the following sections, we will generalize this procedure and extend the use of explicit fields.

2.3 Properties of Explicit Energy Storage Fields

In this section, some properties of energy storage fields in the bond graph representation will be reviewed. For the purpose of generality, the fields in a mixed causality form will be considered. Also, the port variables are allowed to be dependent if the constitutive equations indicate so. We assume that the causality of the field in a model is determined by the SCAP procedure. The causality assignment will depend on its environment and the constitutive equations of the field as well. The following content uses C fields to illustrate the properties of explicit energy storage fields. I fields will have similar properties with appropriate variable representations.

Consider a nonlinear n -port C field with a general form of constitutive equations

$$\Psi_C (e_1, e_2, \dots, e_n, q_1, q_2, \dots, q_n) = \mathbf{0} \quad (2.1)$$

where Ψ_C is a vector of n functions $\psi_{C1}, \psi_{C2}, \dots, \psi_{Cn}$. One can assign the causality in the form shown in Figure 2.4 if and only if the energy stored in the field can be

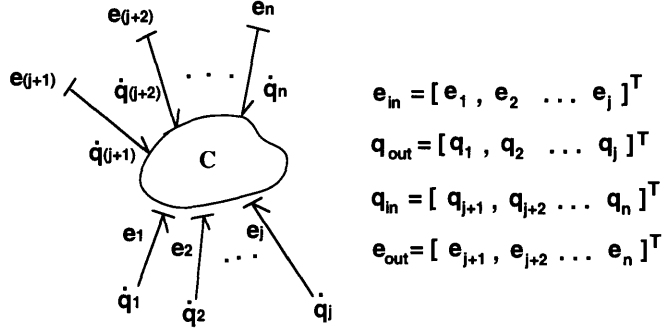


Figure 2.4: A C field with mixed causality assignment.

represented as $E = E(\mathbf{q}_{in}, \mathbf{e}_{in})$, where \mathbf{q}_{in} , \mathbf{e}_{out} are the variables associated with the ports in integral causality and \mathbf{e}_{in} , \mathbf{q}_{out} are the variables associated with the ports in derivative causality. This also implies that the functions in Eqn.(2.1) must be able to be represented in the form

$$\begin{aligned}
 q_1 &= F_1(e_1, e_2, \dots, e_j, q_{(j+1)}, q_{(j+2)}, \dots, q_n) \\
 &\vdots \\
 q_j &= F_j(e_1, e_2, \dots, e_j, q_{(j+1)}, q_{(j+2)}, \dots, q_n) \\
 e_{(j+1)} &= F_{(j+1)}(e_1, e_2, \dots, e_j, q_{(j+1)}, q_{(j+2)}, \dots, q_n) \\
 &\vdots \\
 e_n &= F_n(e_1, e_2, \dots, e_j, q_{(j+1)}, q_{(j+2)}, \dots, q_n)
 \end{aligned} \tag{2.2}$$

According to the principle of thermodynamics [9], the energy of a stable field is a convex function of its extensive variables and a concave function of its intensive variables. In the case of a C field, the generalized displacements q 's are intensive variables and the generalized efforts e 's are extensive variables. Therefore, for a C field, we have the relation

$$\frac{\partial^2 E}{\partial q_i^2} \geq 0 \quad \text{where } i = 1 \dots n \tag{2.3}$$

This in turn implies that the Jacobian matrix formed by the elements $\frac{\partial e_i}{\partial q_j}$ (where $i, j = 1 \dots n$) is positive semi-definite. Also, by Maxwell's reciprocity relations and

Legendre transformations, Eqn (2.2) must satisfy the following relations:

$$\frac{\partial F_k}{\partial e_\ell} = \frac{\partial F_\ell}{\partial e_k} \quad \text{where } k, \ell = 1, 2, \dots, j \quad (2.4)$$

$$\frac{\partial F_{k'}}{\partial q_{\ell'}} = \frac{\partial F_{\ell'}}{\partial q_{k'}} \quad \text{where } k', \ell' = j + 1, j + 2, \dots, n \quad (2.5)$$

$$\frac{\partial F_{k''}}{\partial q_{\ell''}} = -\frac{\partial F_{\ell''}}{\partial e_{k''}} \quad \text{where } k'' = 1, 2, \dots, j \quad (2.6)$$

$$\text{and } \ell'' = j + 1, j + 2, \dots, n$$

The negative sign in Eqn.(2.6) comes naturally from the Legendre transformation due to the cross coupling of the function relation $F_{\ell''}$, $e_{k''}$ and $F_{k''}$, $q_{\ell''}$.

For linear fields, Eqn.(2.2) can be written as

$$\begin{pmatrix} \mathbf{e}_{out} \\ \mathbf{q}_{out} \end{pmatrix} = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{bmatrix} \begin{pmatrix} \mathbf{q}_{in} \\ \mathbf{e}_{in} \end{pmatrix} \quad (2.7)$$

with \mathbf{e}_{out} , \mathbf{q}_{out} , \mathbf{e}_{in} , \mathbf{q}_{in} as defined before. By Eqn.(2.3) to Eqn.(2.7), to form a meaningful energy storage field, this matrix must have at least the following properties:

1. The matrix $\begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{bmatrix}$ is positive semi-definite.
2. \mathbf{C}_{11} and \mathbf{C}_{22} are symmetric positive semi-definite.
3. $\mathbf{C}_{12} = -\mathbf{C}_{21}^T$.

Since the submatrices \mathbf{C}_{11} and \mathbf{C}_{22} are only semi-definite, the causality assignment which can be accepted by the field might be constrained. Also, the independent states which can be contributed by this field might be different from what the causality indicates. By considering the constitutive equations of the field, the following results can be obtained.

Lemma 2.1:

1. *The number of independent states contributed by a multiport C field is determined*

by the rank of the submatrix \mathbf{C}_{11} .

2. If the submatrices \mathbf{C}_{11} and \mathbf{C}_{22} in Eqn.(2.7) are full rank (therefore positive definite), this field can accept any combination of causality assignment if this is not prevented by any other physical constraints, and the number of integral causality indicates the number of independent states.

Furthermore, the rank of the matrices \mathbf{C}_{11} and \mathbf{C}_{22} can be tested by the following causality manipulations.

A proof is given in Appendix A.1.

Lemma 2.2:

1. If the integral causalities on the ports of a field can be reversed, the \mathbf{C}_{11} submatrix of a linear field or the corresponding submatrix in the Jacobian matrix of a nonlinear field must have full rank. If not, the remaining ports with integral causality indicate the rank deficiency in the corresponding submatrix.

2. If the derivative causality on the ports of a field can be reversed, the \mathbf{C}_{22} submatrix or the corresponding submatrix in the Jacobian matrix must have full rank. If not, the remaining ports with derivative causality indicate the rank deficiency in the corresponding submatrix.

A proof is given in Appendix A.1.

Based on the above lemmas, the following rules can be used to determine the number of independent states which can be contributed by a field under the constraints between the field ports.

Proposition 2.1: *In a bond graph model, if a field can accept **reversed** causality on the ports which exhibit integral causality, the number of integral causality will indicate exactly the number of independent states contributed by this field in this model. If not, the number of integral causality remaining indicates the number of type 1 excess states.*

Proposition 2.2: *If a field can accept reversed causality on all ports, this field can accept any combination of causality assignment, and the number of integral causality indicates exactly the number of independent states.*

Proposition 2.3: *If a field can not accept reversed causality on all ports, and some of the ports have integral causalities which are directly imposed by the sources, type 2 excess states might exist. If we keep the integral causality which is imposed by the sources unchanged and reverse the causality of other ports with integral causality, the number of integral causality remaining is the number of type 2 excess states.*

A proof is given in Appendix A.1.

2.4 Model Revision for Structural Property Inspection

In this section, the issues concerning when a model revision is necessary and how it can be done using the explicit field representations are discussed.

2.4.1 The existence of type 1 excess states

As shown in section 2, type 1 excess states cause confusion on the order of dynamic systems. Also, many other structural inspection rules can not be applied properly due to this type of excess states. Therefore, it would be useful if certain test can be performed to detect the existence of type 1 excess states. For this purpose, a testing procedure can be derived [37].

Testing procedure 1 :

- (1) *Apply SCAP procedure.*
- (2) *Remove any energy storage elements with derivative causality and the bonds associated with them.*
- (3) *Remove all causality strokes.*

(4) Relax¹ the causal constraints of the sources and assign derivative causality to the energy storage elements whenever it is possible (under the junction constraints).

(5) If any energy storage element still remains with integral causality, the test fails. That is, correct system order can not be predicted using the SCAP procedure.

Note that step 2 in the above procedure is to remove the constraints which have been taken care of by the SCAP procedure. The purpose of step 3 and 4 is using the derivative causality assignment procedure to expose the constraints which were implicit when using the SCAP procedure as explained at the beginning of this chapter. Although this procedure is easy to implement, removing elements is not an appropriate way of model revision. This is obvious when one would like to use the same model for further structural property inspection. In the following, a procedure using explicit fields is proposed to eliminate the excess constraints without removing essential structural informations.

2.4.2 The use of explicit fields

For the purpose of model revision, we can eliminate the topological constraints (which cause type 1 excess states) by grouping the over-constrained junction structure and the associated energy storage elements into an explicit field. Therefore, the ports of the fields in the revised model would be independent and many pitfalls can be avoided in the further use of the model. The procedure is listed as follows.

Model revision procedure :

(1) Relax the causal constraints of the sources and apply SCAP procedure.

(2) If there is any energy storage element which exhibits derivative causality, identify the energy storage elements of the same type which are directly causally connected² to

¹Here it means to ignore the constraints imposed by the sources and treat them like R elements without any causal constraint.

²Two elements are said to be directly causally connected if there is a causal path between these

this element.

(3) Group the directly casually connected energy storage elements with all the constraints in the junction structure and form an explicit field.

(4) Remove all causality strokes.

(5) Relax the causal constraints of the sources and assign derivative causality to the energy storage elements whenever it is possible (under the junction constraints).

(6) If there is any energy storage element which exhibits integral causality, identify the energy storage elements of the same type which are directly casually connected to this element.

(7) Group the directly casually connected energy storage elements with all the constraints in the junction structure and form an explicit field.

In this procedure, steps 1, 2 and 5, 6 use SCAP and the derivative causality assignment procedure respectively to detect the constraints. In step 3 and step 7, grouping is used to deal with the detected constraints instead of removing elements. By doing so, the model is described by an equivalent yet more appropriate representation. The information concerning dynamic interactions is preserved. Therefore, the revised model can be used for further structural study.

Using this procedure, an explicit field whose port variables are all independent can always be found. According to proposition 2.2, this field accepts any form of causality combination and the number of independent states equals the number of ports which exhibits integral causality. So if the testing procedure is reapplied, the revised model will always pass the test. That is, the topological constraints have been eliminated by this manipulation.

two elements without going through elements other than the junctions.

2.4.3 The existence of type 2 excess states

After the model revision shown in the previous section, the existence of type 2 excess states can be detected by the following procedure.

Testing procedure 2 :

- (1) *Keep the causal constraints of the sources and assign derivative causality to the energy storage elements whenever it is possible (under the junction constraints).*
- (2) *If any energy storage element still remains with integral causality, the system has type 2 excess states. Also, the number of remaining integral causalities indicates the number of excess states.*

Remark : After applying the model revision procedure, all the topological constraints will be eliminated. Therefore, it is not necessary to deal with type 1 excess states. On the other hand, if we skip the model revision procedure and directly apply the testing procedures for type 1 and type 2 excess states, incorrect predictions of excess states might result. The reason is that testing procedure 1 does not leave a complete model for further inspection and testing procedure 2 does not exclusively detect type 2 excess states.

2.5 Applications

In this section, some applications regarding system structural properties are shown to illustrate the use of the explicit field representations described in the previous section.

2.5.1 Independent state variables

Figure 2.5 shows a schematic diagram of a rotor levitated by a set of magnetic bearings and the corresponding bond graph model. In this model, R_s is the electrical resistance; R_e represents the resistance due to eddy currents in the magnetic domain;

C_p represents the reluctance of the permanent magnet and C_{gap} is the energy field between the magnetic domain and the mechanical domain.

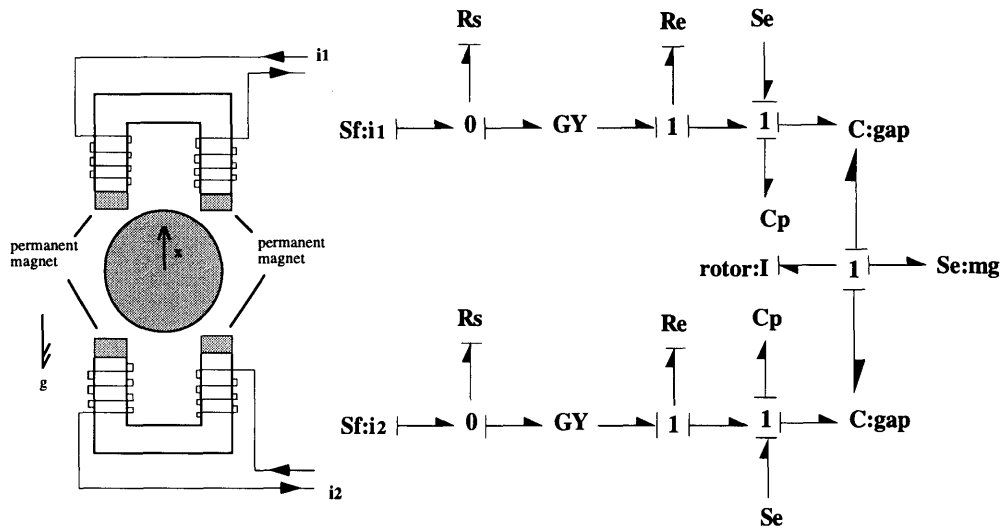


Figure 2.5: Bond graph models of a magnetic bearing.

According to the model revision procedure in the previous section, by assigning derivative causality to energy storage elements, we obtain the direct causal paths as shown in Figure 2.6. Thus an explicit field is obtained by grouping these directly causally related elements as shown in the Figure. Using the port variables of this field, we can obtain a set of independent state equations. Therefore, the number of the independent states is 3. The new model indicates that there are 3 type 1 excess states in the original model of Figure 2.5.

2.5.2 Structural property examination

For the previous example, if one is interested in examining the structural controllability of the system, further causality manipulations are necessary [43]. However, before this test, it can already be concluded that the original model is uncontrollable

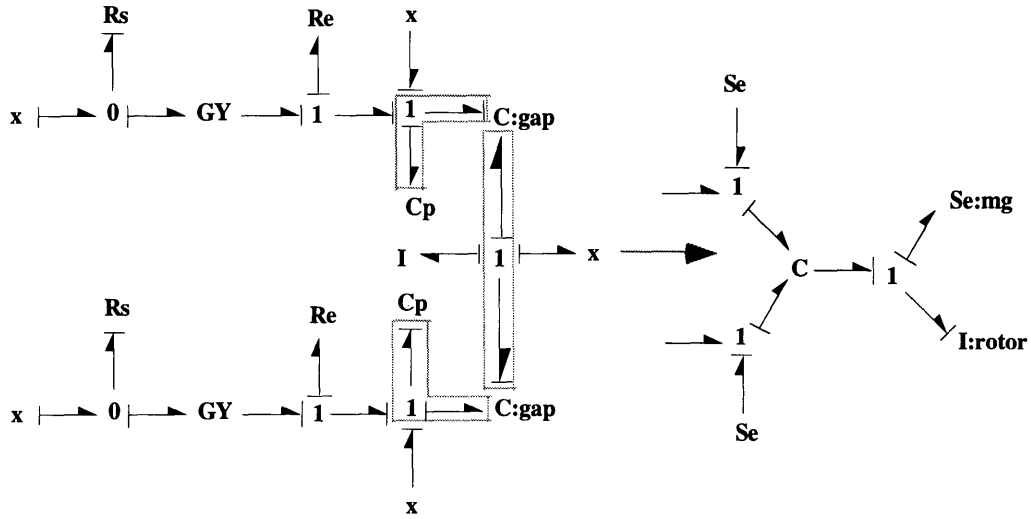


Figure 2.6: Bond graph model with derivative causality assignment and the revised model.

by the result shown in Figure 2.6. Since the derivatives of the states are coupled, it is impossible to achieve full state regulation for the states shown in the original model. On the other hand, using the revised model shown in Figure 2.6, one can find that the new states defined by the field ports are controllable. Note that this is not a special case. Since the excess states can be eliminated by explicit fields, any system which is revised by the model revision procedure in section 4.2 is actually structurally controllable. Therefore, the port variables of the fields contribute a maximum independent set of controllable states associated with the constrained energy storage elements. If one really would like to control the states in the original system, either some dynamic elements or even sources must be added to break the topological dependency. This provides a guidance of designing the physical systems.

2.5.3 The equilibrium states

From the bond graph model, the equilibrium states of the system can be directly obtained if the causalities are used properly [8, 13]. At equilibrium points, the derivatives

of the states are identically zero. This status can be represented in the bond graph by setting the flow to a C element and the effort to an I element zero. Thus if we replace the C nodes by zero flow sources and I nodes by zero effort sources and propagate the causality, the equilibrium states (the effort to the C nodes and the flow to the I nodes) then can be solved directly from the graph. It is found that this procedure can only be applied to the revised model. Using the original model, causality conflicts would result if the sources mentioned above were imposed. As shown in Figure 2.7, the equilibrium states are $e_1 = n \cdot i_1$; $e_2 = n \cdot i_2$; $e_3 = mg$; $f_4 = 0$.

This result shows that in the equilibrium status, the magnetomotive force on port 1 and 2 are $n \cdot i_1$ and $n \cdot i_2$ respectively; the force acting on the rotor equals to the gravity force mg ; and the velocity of the rotor is zero.

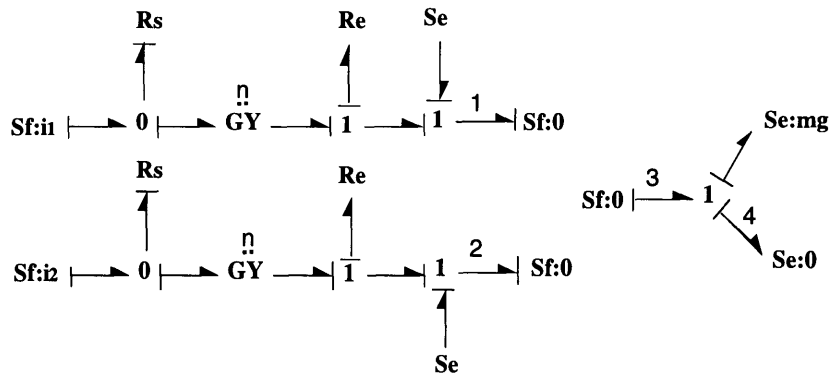


Figure 2.7: Equilibrium states from bond graph model with field representation.

2.6 Computer Implementation

After working out several examples, one might find that usually the C elements around 1 junction and the I elements around 0 junction are likely to generate excess states. Thus they should be grouped into an explicit field. For example, the capacitances

in Figure 2.8 (a) should be viewed as an equivalent capacitance in Figure 2.8 (b) so that correct structure properties can be determined. However, this kind of rule is too ambiguous for system property inspection, especially for computer implementation. To develop an application program for automated model processing, the elements and bonds which should be included in a field need to be exactly identified. This is an important reason of adopting the causal searching rules in this thesis. As shown in Figure 2.6, the part of the system which belongs to a local explicit field is identified without any ambiguity. The computer program may then blindly obtain a 6 port field as shown in Figure 2.9.

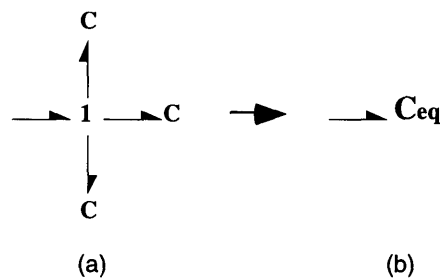


Figure 2.8: A simple example of grouping capacitances.

By proposition 2.1 in Section 2.3, this is a legitimate representation since the 3 integral causality indicates that there are actually only 3 independent states contributed by the field. However, by examining the equations of this field, an algorithm can be easily derived to release the constraints between the ports. For example, the

where

$$[\mathbf{JS}] = \left[\begin{array}{cccccc|cccccc} 0 & -1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ - & - & - & - & - & - & - & - & - & - & - & - \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

Because there are only 3 independent columns in the left lower submatrix in the above equation, it is clear that the equations can be rewritten as

$$\begin{pmatrix} e'_1 \\ e'_7 \\ e'_{11} \\ - \\ f_3 \\ f_4 \\ f_5 \\ f_8 \\ f_9 \\ f_{10} \end{pmatrix} = \left[\begin{array}{cccc|cccc} 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ - & - & - & - & - & - & - & - & - \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right] \begin{pmatrix} f_1 \\ f_7 \\ f_{11} \\ - \\ e_3 \\ e_4 \\ e_5 \\ e_8 \\ e_9 \\ e_{10} \end{pmatrix}$$

and 3 extra 1 junctions can be defined. Thus a 3-port C field in Figure 2.6 can be obtained automatically. The details of this process are shown in Appendix A.2.

2.7 Conclusion

In this chapter, a method using the explicit field representations is proposed to eliminate the excess states for the purpose of structure inspections. For the constraints caused by linear junctions, it was shown that the excess states caused by the topology structures can be totally eliminated by field representations. The excess states caused by the imposed sources then can be identified properly. Several applications of the explicit fields are shown to illustrate their use.

Chapter 3

Zero Dynamics of Physical Systems from Bond Graph Models - Part I : SISO Systems

3.1 Introduction

Zero dynamics is an important feature in system analysis and controller design. Its behavior plays a major role in determining the performance limits of certain feedback systems. For example, it is known that perfect tracking of arbitrary trajectory can not be achieved by any controller if the zero dynamics is unstable [39]; H_∞ controller design needs the sensitivity functions to satisfy certain interpolation condition at the location of non-minimum phase zeros [10]. Thus the bounds of the achievable performance is partially determined by the zero dynamics; The application of input-output linearization schemes also requires the stability of zero dynamics [21]. Since the intrinsic zero dynamics can not be influenced by feedback compensation, it is important to design physical systems so that they possess desired zero dynamics. However, the derivation of the zero dynamics is usually complicated, especially if a form which is closely related to the physical system and suitable for design is required. In order to address such issues, a method of designing the zero dynamics from the bond graph point of view is proposed in this chapter.

The zero dynamics is interpreted as the resultant internal dynamics when suitable initial conditions and control inputs are applied to maintain the outputs zero for all time. For simple physical systems, part of or full zero dynamics may be obtained by

direct inspection of the system structure. For example, some heuristic rules are found to identify the zeros for certain class of linear SISO mechanical systems [31]. This can be done directly from the system models consisting of a series of mass-spring-damper units. The idea is extended in [47] by using the bond graph representation. For a class of nonlinear SISO systems, the zero dynamics can be obtained by recognizing the junction structure patterns. These results point out a potential direction of structural analysis for zero dynamics directly from model representations.

Bond graph models describe the dynamic behavior of physical systems by the connection of idealized lumped elements based on the principle of energy conservation [22]. These network-like models provide very useful insights to the structure of dynamic systems. With the help of the encoded structures and the causality technique, the structural inspection and the derivation of zero dynamics can be generalized for a much broader class of systems. In this chapter, the definition of zero dynamics in the differential geometric approach is incorporated with the causality manipulation in the bond graph representation. By doing so, the state equations of zero dynamics can be easily obtained. The system elements which are responsible for the zero dynamics can be identified. In addition, if isolated subsystems which exhibit the zero dynamics exist, they can be found. These subsystems, which describe the zero dynamics, are often of low order and consequently are easier to analyze. Thus, the design of physical systems including the consideration of zero dynamics become straightforward. Since this approach does not depend on the heuristic rules for systems with specific patterns, the results can be applied for general SISO systems and can be generalized for MIMO systems.

In section 3.2, the definition of the zero dynamics in the differential geometric approach is reviewed. Section 3.3 describes the general procedure of deriving the zero dynamics of SISO systems from the bond graph models. Section 3.4 shows some

applications of the proposed procedure, and the conclusion is given in Section 3.5.

3.2 Zero Dynamics in the Differential Geometric Approach

In this section, the definition of the zero dynamics from the differential geometric point of view is reviewed [21, 39]. This approach deals with a class of nonlinear systems with the general form

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})\mathbf{u} \quad (3.1)$$

$$\mathbf{y} = \mathbf{h}(\mathbf{x}) \quad (3.2)$$

where $\mathbf{u} \in R^p$ is the input vector, $\mathbf{y} \in R^m$ is the output vector and $\mathbf{x} \in R^n$ represents the state vector. The zero dynamics is defined as the internal dynamics of the system when the required initial conditions and controls are applied to keep the outputs zero for all time. Thus the zero dynamics can be described as

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})\mathbf{u}^* \quad (3.3)$$

with suitable initial states. The conditions for \mathbf{u}^* to exist and how the system evolves under these conditions can be derived in a rather systematic and rigorous way by considering a local coordinate transformation problem. This analysis will in turn lead to the application of input-output linearization and certain dynamic decoupling problems. In this chapter, the relevant existing results are stated since the details can be found in [21, 39].

From the differential geometric approach, it can be derived that the zero dynamics of an SISO system in the form of Eqn.(3.1), (3.2) will evolve on the subset

$$Z^* = \{ \mathbf{x} \in R^n : h(\mathbf{x}) = \dots = L_{\mathbf{f}}^{r-1}h(\mathbf{x}) = 0 \} \quad (3.4)$$

$$\text{or equivalently, } Z^* = \{ \mathbf{x} \in R^n : y(\mathbf{x}) = \dot{y}(\mathbf{x}) \dots = y^{(r-1)}(\mathbf{x}) = 0 \} \quad (3.5)$$

when the system is under the control $u^*(\mathbf{x})$ given by,

$$u^*(\mathbf{x}) = \frac{-L_{\mathbf{f}}^r h(\mathbf{x})}{L_g L_{\mathbf{f}}^{r-1} h(\mathbf{x})} \quad (3.6)$$

and the initial conditions lie in the subset described by Eqn.(3.4). In the above equations, the symbol L represents the Lie derivative which operates as $L_{\mathbf{a}} b(\mathbf{x}) = \nabla b \cdot \mathbf{a}$; r represents the relative degree of the system; $L_{\mathbf{f}}^r h(\mathbf{x})$ represents an r^{th} order consecutive Lie derivative $L_{\mathbf{f}} \dots L_{\mathbf{f}} h(\mathbf{x})$. An SISO system is said to have a relative degree r at a point \mathbf{x}^o if [21]

- (i) $L_g L_{\mathbf{f}}^k h(\mathbf{x}) = 0$ for all \mathbf{x} in a neighborhood of \mathbf{x}^o and all $k \leq r - 1$.
- (ii) $L_g L_{\mathbf{f}}^{r-1} h(\mathbf{x}^o) \neq 0$.

Thus, assuming that the control of Eqn.(3.6) is applied, the zero dynamics can be viewed as the residual dynamics of Eqn.(3.1) under the constraints in Eqn.(3.4). By substituting these constraint equations into Eqn.(3.1), a minimal set of state equations which represents the zero dynamics can be obtained.

Note that the representation of the zero dynamics is not unique. If a local coordinate transformation $\mathbf{z} = \Phi(\mathbf{x})$ is found to transform the state equations into the normal form,

$$\begin{aligned} \dot{z}_1 &= z_2 \\ \dot{z}_2 &= z_3 \\ &\dots \\ \dot{z}_{r-1} &= z_r \end{aligned} \quad (3.7)$$

$$\begin{aligned} \dot{z}_r &= b(\xi, \eta) + a(\xi, \eta)u \\ \dot{\eta} &= q(\xi, \eta) \\ y &= z_1 \end{aligned} \quad (3.8)$$

where

$$\xi = \begin{bmatrix} z_1 \\ \dots \\ z_r \end{bmatrix}, \quad \eta = \begin{bmatrix} z_{r+1} \\ \dots \\ z_n \end{bmatrix} \quad (3.9)$$

then the zero dynamics can be easily identified as

$$\dot{\eta} = q(0, \eta) \quad (3.10)$$

by choosing $\xi(0) = \mathbf{0}$ and $u = -\frac{b(\xi, \eta)}{a(\xi, \eta)}$. However, the calculation of this transformation usually requires more effort than the derivation from the original equations. Also, the coefficients in the normal form often do not have any physical meaning. Thus, Eqn.(3.1) and Eqn.(3.4) are more suitable for our design purpose.

3.3 Zero Dynamics of SISO Systems from Bond Graph Models

The bond graph representation encodes the structure of dynamic systems into abstract graph symbols. Thus the models can provide a useful guidance for the derivation of the zero dynamics, which is far from straightforward as shown in the previous section. For certain systems, the properties of zero dynamics can be found simply by inspection. In this section, we will discuss the derivation and inspection of zero dynamics with the help of bond graph models for SISO systems. The extension to MIMO systems follows directly from this method as will be shown in the next chapter.

A key technique in using the bond graph models is the use of causal implications. The causality indicates the dependency of bond variables ¹in the bond graph model. Thus, it in turn implies the way of deriving the equations. For example, the state equations which describe the system dynamics can be derived systematically by the use of causality [36]. From this procedure, it can be shown that, in general², the

¹The effort or flow variable associated with a bond is called the bond variable.

²If the system contains excess states, this statement is not true [17].

integral causality in the bond graph models indicates an independent state variable. By the help of causality, many useful system structural properties can be identified [22, 24, 41, 43, 45]. In what follows, a systematic way of identifying the independent states for the zero dynamics will be presented. Then the constraints of Eqn.(3.4) are imposed by certain causality manipulations. Finally, the partitions from the original dynamics which exhibit the zero dynamics appears directly from the inspection of causality.

3.3.1 Definitions

In this chapter, the considered bond graph models contain³ one port dissipative element R ; one port energy storage elements C , and I ; linear junctions 1 , 0 , TF , GY ; and sources S_e , S_f . To facilitate the description of model processing, the following definitions are introduced.

Causal path : In a bond graph model, a series of bond variables which connects one specific variable to another according to the causality assignment is called a causal path between these two variables.

Shortest causal path : Among the alternative paths connecting two bond variables, the one that yields the minimum number of independent energy storage elements⁴ less dependent energy storage elements⁵ on the path is called the shortest causal path.

Simple shortest causal path : A shortest causal path which passes through no dependent energy storage elements is called a simple shortest causal path.

Causal input (output) variable : The bond variable which represents an independent (dependent) variable in the constitutive relation of a bond graph element

³For the models which contain modulated junctions, fields or even subsystem representations, procedures similar to the one described in this chapter, although more complicated, can be developed and used in the same manner.

⁴Elements with integral causalities.

⁵Elements with derivative causalities.

according to the causality assignment is called a causal input (output) variable.

For simplicity, the shortest causal path mentioned in this chapter means simple shortest causal path unless specifically emphasized.

3.3.2 Relative degree

Since our purpose of modeling is for the control or dynamic analysis of physical systems, the output of interest is mostly related to only one of the states like displacement, velocity, and not a combination of the states. In this case, the output is related to certain bond variable in the graph. Thus a causal path from the input to the output variable can always be found. This kind of causal path might not be unique. With the use of these causal paths, the following propositions are derived to identify the relative degree of SISO systems.

Proposition 3.1: *In an SISO bond graph model with the choice of suitable causality assignment, a simple shortest causal path which connects the input to the output variable (a state variable) always exists.*

Proposition 3.2: *In an SISO bond graph model, the number of elements with integral causality on the simple shortest causal path described in proposition 3.1 indicates the structural relative degree of this system.*

A proof is given in Appendix B.1

Note that proposition 3.1 rules out the necessity of dealing with the shortest causal path which contains derivative causalities. As will be shown in the following sections, this provides a clear choice of candidate states for the zero dynamics. Proposition 3.2 is a modified version of the rules which are derived in [30, 47].

The above propositions follow from the interpretation of the relative degree as the number of times one has to differentiate the output so that the input explicitly appears. Since the relative degree is also the difference between the order of the overall system dynamics and the order of the zero dynamics, the result also implies

that the number of integral causalities which are not on the shortest path indicates the order of zero dynamics. In fact, each independent energy storage element which is not on this path contributes an independent state for the zero dynamics. Using these states, a set of state equations representing the zero dynamics can be derived.

Example: To illustrate the use of these propositions, a simple mass-damper-spring unit shown in Figure 3.1 is considered as an example. In this system, the input is the force applied on mass 1 and the output is chosen as the velocity of mass 3. A bond graph model of this system is shown in Figure 3.2. According to the definition of causal path, there could be several choices. For example, Figure 3.3 and Figure 3.4 show different choices of causal paths. These two causal paths pass 5 and 3 energy storage elements respectively. Obviously, the causal path in Figure 3.4 is the shortest causal path. According to the propositions, this system has a relative degree 3.

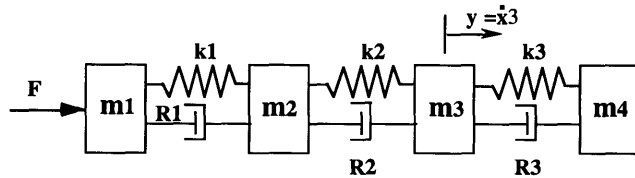


Figure 3.1: A mass-damper-spring unit.

3.3.3 Zero dynamics

By the definition in section 2, the zero dynamics evolve on the subset described by Eqn.(3.4). On the bond graph model, if we trace back the shortest causal path from the output $y = h(x)$ to the input u , it will be found (See Appendix B.2) that the causal output variable of each energy storage element (the state variable) on this path appears explicitly in order when the 1^{st} to $(r - 1)^{th}$ derivatives of the output

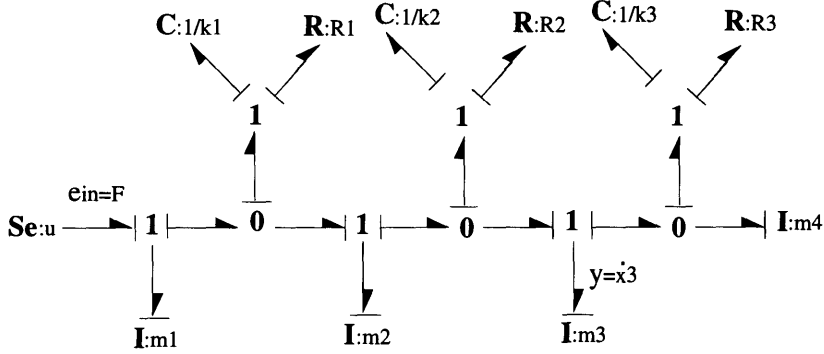


Figure 3.2: A bond graph model of the mass-damper-spring unit.

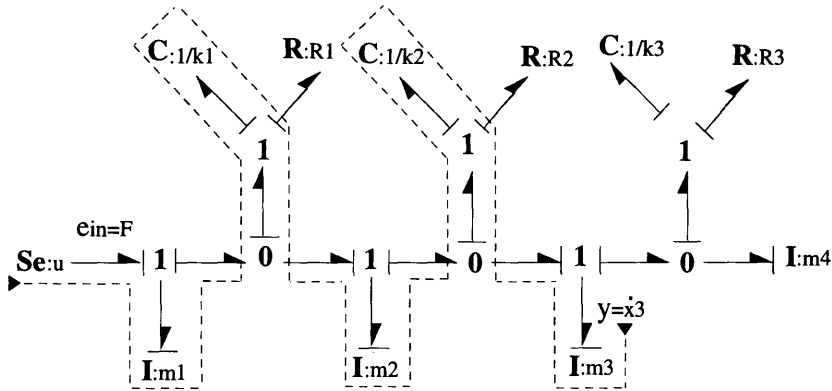


Figure 3.3: A choice of causal path.

are taken. Taking the model in Figure 3.4 as an example⁶, the output is $y = f_3$. The causal output variable of m_2 , which is f_2 , first appears in the expressions of \dot{y} as shown in Eqn.(3.19). The causal output variable of m_1 , which is f_1 , first appears in the expression of \ddot{y} and not y or \dot{y} as shown in Eqn.(3.20). Finally, the input u appears when the third derivative of the output is taken as shown in Eqn.(3.21).. This is made clear by the meaning of the shortest causal path itself. Since $L_f^k h(\mathbf{x})$ equals the k^{th} derivative of y when $k \leq r - 1$, each algebraic equation in Eqn.(3.4) indicates that

⁶The state equations are derived in the following section.

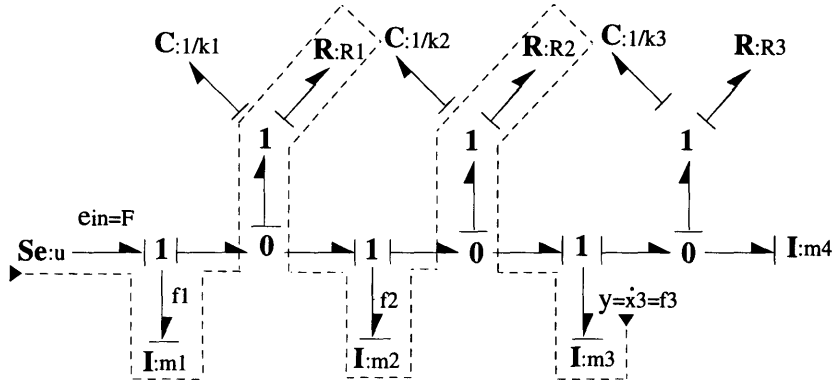


Figure 3.4: The shortest causal path.

exactly one causal output variable of the energy storage element on this path becomes dependent to other states. This dependency of bond variables can be represented by the causality. Thus the zero dynamics can be derived by the following procedure. We will refer to this procedure as the Zero Dynamics Identification Procedure (ZDIP) in this thesis.

ZDIP procedure

- (1) *Apply the normal Sequential Causality Assignment Procedure (SCAP) to the system bond graph model.*
- (2) *Determine the shortest causal path between the input and the output as described in proposition 3.1. If there are more than one of such causal paths, pick one of them arbitrarily⁷.*
- (3) *Remove the bond which is related to the output variable, and assign a zero value to the junction which this bond is connected to. This indicates that the common effort of the 0 junction or the common flow of the 1 junction is zero.*
- (4) *Reverse the causality on the bond which is connected to the input.*

⁷This situation indicates that alternative choices of independent state variables for the zero dynamics representation are possible.

- (5) *Replace every integral causality on this causal path by a derivative causality.*
- (6) *Keep every integral causality which is not on this causal path unchanged. These are the candidates of the independent state variables for the zero dynamics.*
- (7) *Complete the causality assignment while imposing the junction constraints.*
- (8) *Derive the zero dynamics according to the causality assignment.*
- (9) *If the above derivation requires the causal outputs from the elements whose causality has been reversed, solve for these variables using the algebraic equations in Eqn.(3.4) so that they can be represented by the independent state variables determined in step 6.*

Note that step 2 is to identify the relative degree and determine which variables can be state variable candidates of the zero dynamics. Step 3 imposes the algebraic equations $y = h(x) = 0$ and $\dot{y} = L_{\mathbf{f}}h(x) = 0$. Step 4 reflects the fact that when deriving the zero dynamics, the input u^* is determined by the system states. Steps 5, 6 and 7 impose the rest of the algebraic equations in Eqn.(3.4) by setting the effort of the inertance and the flow of the capacitance on the shortest causal path as the causal output from the energy storage elements (please see Appendix B.2 for details). In this case, the causality is used to indicate the dependency of the bond variables. The constitutive equations of the corresponding energy storage elements will not be used in this zero dynamics model. Instead, the constraints in Eqn.(3.4) are used to determine the causal outputs. Using the equations $L_{\mathbf{f}}^k h(\mathbf{x}) = 0$, where $0 \leq k \leq r - 1$, these efforts and flows can be written as functions of the independent state variables in step 6. If necessary, these equations will be used in step 9. However, step 9 is only necessary for certain complicated systems where isolated subsystems which represent the zero dynamics do not exist. For simple systems, this step can usually be skipped.

Using this procedure, the zero dynamics of a nonlinear SISO system can be derived systematically. The physical elements or subsystems which are responsible for the zero

dynamics can be easily identified. Thus, the properties such as the stability of zero dynamics can be studied from the graph. By doing so, the dynamic features can be related directly to the physical system configuration. This would be very useful in the design of physical systems.

Example: Consider the system of Figure 3.1, the zero dynamics of this system can be easily identified by the ZDIP procedure. Since the shortest causal path is the one shown in Figure 3.4, the energy storage elements I_1 and I_2 are on the path. Also, because the output is the velocity of element I_3 , the bond associated with the element I_3 is removed according to step 3. For step 4, the causality of the effort source is reversed. For step 5, causalities associated with the elements I_1 and I_2 are reversed. After the causality assignment is complete, the model is shown in Figure 3.5. The dynamics represented by this model is the zero dynamics of the original model.

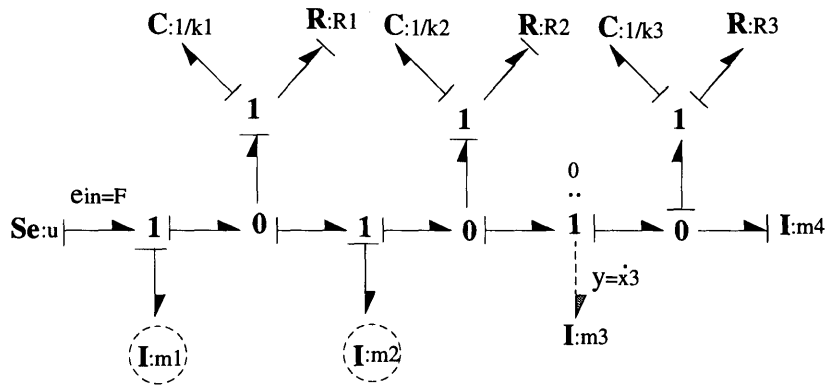


Figure 3.5: The zero dynamics model.

To illustrate the purpose of these steps in the ZDIP procedure, the state equations of this system are derived as follows. For simplicity, all elements are assumed linear. The velocity associated with the masses m_1 , m_2 , m_3 and m_4 are f_1 , f_2 , f_3 and f_4 . The forces associated with the springs C_1 , C_2 and C_3 are e_1 , e_2 and e_3 .

$$\dot{f}_4 = \frac{1}{m_4}[e_3 - (f_3 - f_4)R_3] \quad (3.11)$$

$$\dot{e}_3 = (f_3 - f_4)k_3 \quad (3.12)$$

$$\dot{f}_3 = \frac{1}{m_3}[e_2 + (f_2 - f_3)R_2 - e_3 - (f_3 - f_4)R_3] \quad (3.13)$$

$$\dot{e}_2 = (f_2 - f_3)k_2 \quad (3.14)$$

$$\dot{f}_2 = \frac{1}{m_2}[e_1 + (f_1 - f_2)R_1 - e_2 - (f_2 - f_3)R_2] \quad (3.15)$$

$$\dot{e}_1 = (f_1 - f_2)k_1 \quad (3.16)$$

$$\dot{f}_1 = \frac{1}{m_1}[u - e_1 - (f_1 - f_2)R_1] \quad (3.17)$$

According to the differential geometric approach, the zero dynamics will evolve on the set defined by Eqn (3.4). In other words, if the constraints in Eqn (3.4) are imposed on the system dynamics, the residual dynamics will be the zero dynamics. In this case, since the relative degree has been found as 3 in the previous section, the constraints are $y^{(k)} = L_f^k h(\mathbf{x}) = 0$, where $0 \leq k \leq 2$.

$$y = f_3 = 0 \quad (3.18)$$

$$\dot{y} = \dot{f}_3 = \frac{1}{m_3}[e_2 + (f_2 - f_3)R_2 - e_3 - (f_3 - f_4)R_3] = 0 \quad (3.19)$$

$$\ddot{y} = \ddot{f}_3 = \frac{1}{m_3}[\dot{e}_2 + (\dot{f}_2 - \dot{f}_3)R_2 - \dot{e}_3 - (\dot{f}_3 - \dot{f}_4)R_3] = 0 \quad (3.20)$$

Suppose these constraints are substituted into the state equations, the residual dynamics should be of order 4. Therefore, there will be only 4 independent states in the residual dynamics. The first step to do the substitution is to determine these independent states. Recall that the relative degree can be determined by the shortest causal path. Since the relative degree is the difference between the order of the overall system and the zero dynamics, it is reasonable to choose the state variables which are

not on the shortest causal path as the independent states. Therefore, the purpose of the substitution is obviously to eliminate the state variables which are on the shortest causal path. In this case, they are f_3 , f_2 and f_1 .

First of all, $y = f_3 = 0$ can be easily substituted into Eqn.'s (3.11) and (3.12). Also, $\dot{y} = \dot{f}_3 = 0$ eliminates Eqn. (3.13). These substitutions are shown in the bond graph of Figure 3.5 by removing the bond which is associated with the output in step 3 of the ZDIP procedure. Eqn.(3.14) should be preserved since e_2 is not on the shortest causal path. However, the right hand side of this equation contains state variables f_2 and f_3 , which should be eliminated. Using Eqn. (3.19), $(f_2 - f_3)$ can be represented as $\frac{1}{R_2}[-e_2 + e_3 + (f_3 - f_4)R_3]$. This is exactly what the causality shows in the zero dynamics model of Figure 3.5. Finally, using Eqn.(3.20), \dot{f}_2 can be written as $\frac{1}{R_2}[-\dot{e}_2 + \dot{f}_3 R_2 + \dot{e}_3 + (\dot{f}_3 - \dot{f}_4)R_3]$, where $\dot{f}_3 = 0$ from Eqn. (3.19), \dot{f}_4 can be found in Eqn. (3.11), and \dot{e}_2 has been derived above. In the bond graph model, this is shown by reversing the causality of the bond associated with I_2 . Therefore, Eqn. (3.15) now can be removed. Also, $(f_2 - f_3)$ can be represented as $\frac{1}{R_1}[m_2 \dot{f}_2 - e_1 + e_2 - (f_2 - f_3)R_2]$, where $(f_2 - f_3)$ and \dot{f}_2 have been derived above. This substitution is again exactly what the causality shows in the zero dynamics model. Up to this point, the zero dynamics has been derived. However, if one is interested in finding the necessary control input for keeping the output zero for all time, Eqn. (3.21) will be considered.

$$y^{(3)} = f_3^{(3)} = \frac{1}{m_3}[\ddot{e}_2 + (\ddot{f}_2 - \ddot{f}_3)R_2 - \ddot{e}_3 - (\ddot{f}_3 - \ddot{f}_4)R_3] = 0 \quad (3.21)$$

In this equation, \dot{f}_1 will appear in the expression of \ddot{f}_2 . Using Eqn.(3.17), the necessary input u^* can be found as $m_1 \dot{f}_1 + e_1 + (f_1 - f_2)R_1$, where \dot{f}_1 can be found using Eqn. (3.21), and $(f_1 - f_2)$ has been derived earlier. This derivation is shown in the bond graph of Figure 3.5 by reversing the causality associated with the effort source.

By examining the causality in the model of Figure 3.5, the zero dynamics of

this linear(nonlinear) system can be partitioned as shown in Figure 3.6. Note that in the partitioned model, the elements in each individual subsystem are energetically coupled. Between the subsystems, there are dashed lines with arrows, which represent one way interactions. For example, the input of subsystem 2 is from the output effort of subsystem 1. On the other hand, subsystem 1 is isolated, since the input flow to this subsystem is zero. The directed line associated with I_2 indicates that the input to subsystem 3 depends on not only the output from subsystem 1, but also on the expression of \dot{f}_2 . From the above derivation, it is clear that \dot{f}_2 can be represented by a function of e_2 , e_3 and f_4 . If the system is linear, these one way interactions have no contribution to the eigenvalues of the zero dynamics. Therefore, the zeros will be the eigenvalues of the individual subsystems: $\frac{k_1}{R_1}$, $\frac{k_2}{R_2}$, and the roots of $m_4s^2 + R_3s + k_3 = 0$. If the system is nonlinear, the one way interactions should be considered when determining the stability of the zero dynamics.

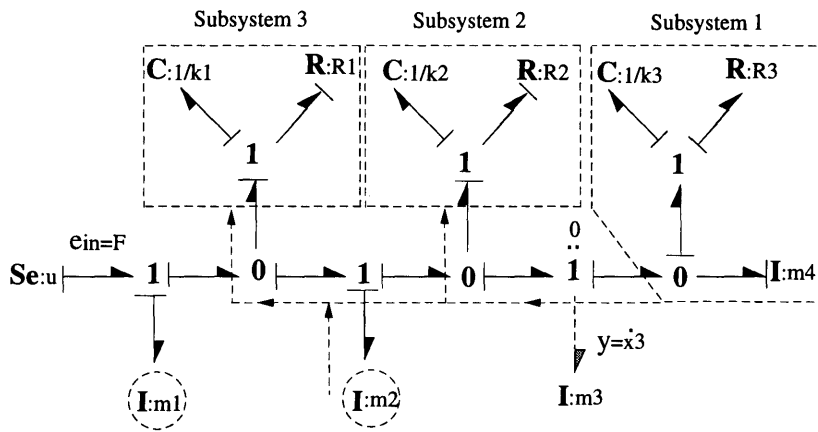


Figure 3.6: The partitioned zero dynamics model.

3.4 Applications

As shown in the previous section, the complicated derivations can be represented by simple causality manipulations. Therefore, using the ZDIP procedure, the zero dynamics can be identified by simple causality reasoning. The following applications illustrate the use this procedure.

3.4.1 Systems with the input and the output on the same bond

If the input and the output variables are on the same bond, we can find that the causal path from the input to the output will go through at most one energy storage element. Therefore, by the rules in section 3.2, the relative degree ⁸ is either 0 or 1. In either case, steps 1, 2, 3 and 5 in the ZDIP procedure can be skipped. So after the source-imposed causality is reversed, the model would represent the zero dynamics. If the system is linear, the eigenvalues of this dynamics would be the zeros of the original system. For example, the linear system shown in Figure 3.7 (a) has the state equation as

$$\begin{aligned}\frac{de}{dt} &= \frac{1}{C}f = \frac{1}{C}(f_{out} - f_1) = \frac{1}{C}\left(\frac{e_{in} - e}{R_2} - \frac{e}{R_1}\right) \\ &= \frac{-1}{C} \frac{R_1 + R_2}{R_1 R_2} e + \frac{1}{C R_2} e_{in}\end{aligned}\quad (3.22)$$

, thus the pole is at $\frac{-1}{C} \frac{R_1 + R_2}{R_1 R_2}$. Using the ZDIP procedure, the zero dynamics is given by the system shown in Figure 3.7 (b). According to the causality assignment, since the source of this bond graph is $f_{out} = 0$, the state equation can be derived as

$$\frac{de}{dt} = \frac{1}{C}f = \frac{1}{C}\left(-\frac{e}{R_1}\right)\quad (3.23)$$

This indicates that the original system has a zero at $\frac{-1}{C R_1}$.

⁸This result is consistent with the well-known network theory.

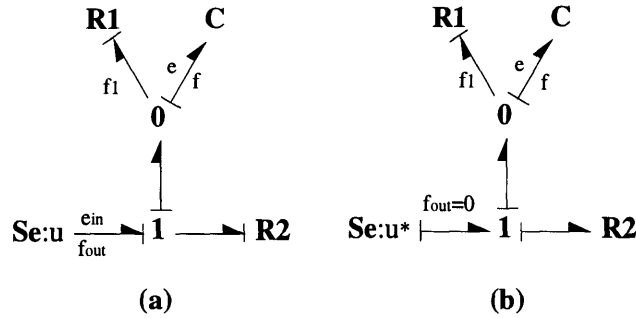


Figure 3.7: (a) Bond graph model of a simple RC system. (b) The zero dynamics.

3.4.2 Systems with simple structures

For certain simple systems where step 9 can be skipped, the zero dynamics usually turn out to be in a special chain form,

$$\begin{aligned} \dot{\mathbf{x}}_1 &= \mathbf{f}_1(\mathbf{x}_1) \\ \dot{\mathbf{x}}_2 &= \mathbf{f}_2(\mathbf{x}_1, \mathbf{x}_2) \\ &\dots \\ \dot{\mathbf{x}}_{n-r} &= \mathbf{f}_{n-r}(\mathbf{x}_1, \dots, \mathbf{x}_{n-r}) \end{aligned}$$

where $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n-r}$ are the state vectors of the subsystems partitioned by the ZDIP procedure. If all these subsystems are globally stable, then the stability of the overall zero dynamics can be determined by the stability of the isolated subsystems,

$$\begin{aligned} \dot{\mathbf{x}}_1 &= \mathbf{f}_1(\mathbf{x}_1) \\ \dot{\mathbf{x}}_2 &= \mathbf{f}_2(\mathbf{x}_2) \\ &\dots \\ \dot{\mathbf{x}}_{n-r} &= \mathbf{f}_{n-r}(\mathbf{x}_{n-r}) \end{aligned}$$

Otherwise, the range of stability must be defined. This will influence the acceptable initial condition range in the controller design of nonlinear systems.

One example of such systems is the model of Figure 3.6 in the previous section. In that model, the zero dynamics is derived for a simple mechanical system. However, the systems of this class do not have to be in the same energy domain. Since the interactions between the subsystems are basically determined by the system structure, a chain form can exist even if the system contains several energy transductions. For example, the model shown in Figure 3.8 also has zero dynamics in such a chain form. This is a simplified dynamic model of a speaker. The output is selected as the velocity v of the inertance on the right hand side of the transformer. The dotted line in Figure 3.8 shows the shortest causal path. The causality assignment in determining the zero dynamics is shown in Figure 3.9. By examining the causality, it is found that the zero dynamics can be partitioned as shown in Figure 3.10. Note that the zero denoted on the 0 junction indicates that the common effort is zero. Thus the input to subsystem 1 is zero. On the other hand, the input to subsystem 2 is the state from subsystem 1. So the zero dynamics is in a chain form. If the system is linear, it will have a zero at $\frac{-1}{C_2} \frac{R_1 + R_2}{R_1 R_2}$, which is determined by subsystem 1 and a zero at the origin, which is determined by subsystem 2.

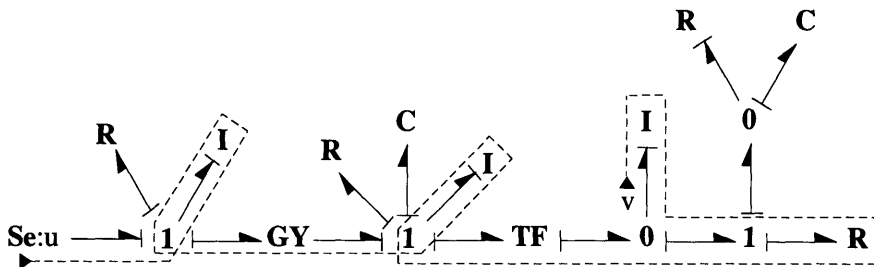


Figure 3.8: Bond graph model of a speaker.

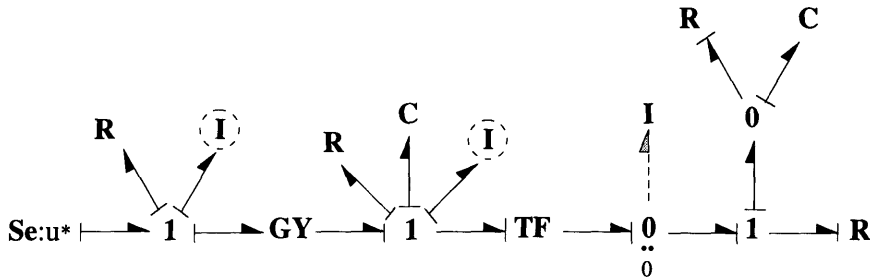


Figure 3.9: The causality assignment after ZDIP procedure.

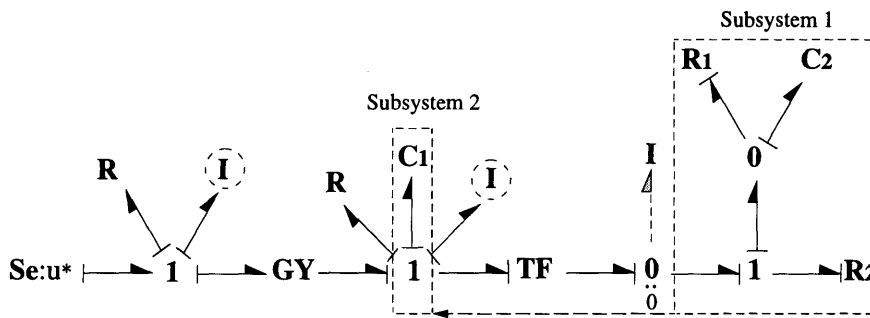


Figure 3.10: The zero dynamics identified by ZDIP procedure.

3.4.3 A Design Example : Prosthesis Arm Design

As shown in the above examples, the ZDIP procedure can be used for a quick analysis of dynamic models. However, a more important purpose of finding the zero dynamics is to facilitate system design. In the following, a design example is shown to demonstrate that the ZDIP procedure allows the system designer to conceive a system configuration with desired zero dynamics.

A schematic of a simple prosthesis mechanism [1] is shown in Figure 3.11. The input of this system is the current to the motor and the output is the angular velocity of the arm prosthesis. The corresponding bond graph model is shown in Figure 3.12. Note that due to the kinematic relations, the model contains a junction loop. By

applying the ZDIP procedure, the zero dynamics model is shown in Figure 3.13. Assuming that all elements are linear, the following equations can be derived by the

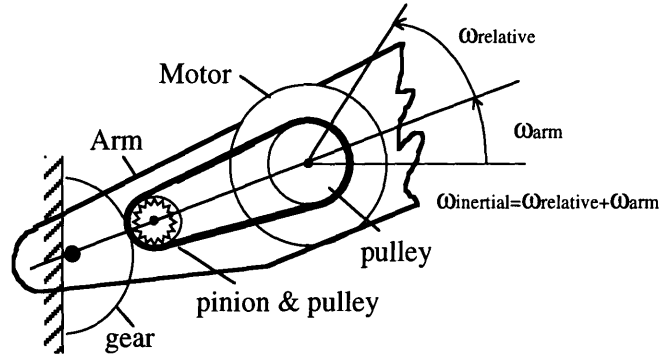


Figure 3.11: The schematic of an arm prosthesis.

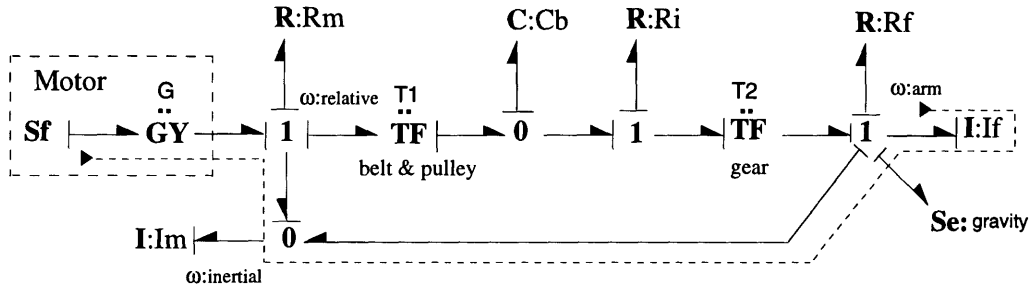


Figure 3.12: The bond graph model of an arm prosthesis.

causality shown in Figure 3.13,

$$\dot{p}_m = \frac{q_C}{C_b} T_2 \quad (3.24)$$

$$\dot{q}_C = \frac{f_s}{T_1} \quad (3.25)$$

From the original model, f_s can be found as $\frac{p_m}{I_m}$ ($\omega_{\text{arm}} = 0$). Thus the zero dynamics is

$$\dot{p}_m = \frac{q_C}{C_b} T_2 \quad (3.26)$$

$$\dot{q}_C = \frac{p_m}{I_m T_1} \quad (3.27)$$

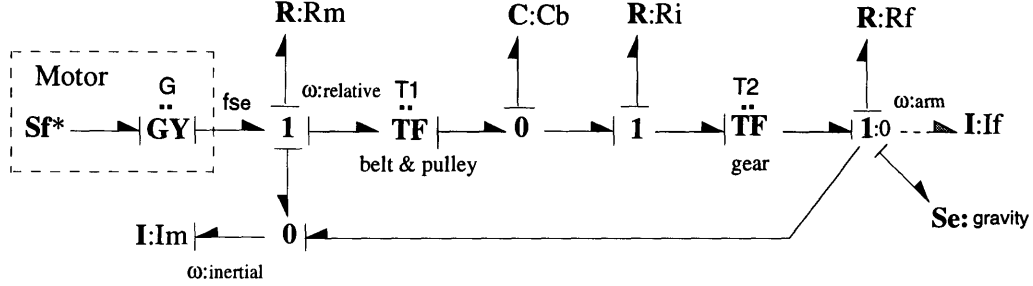


Figure 3.13: The zero dynamics model of an arm prosthesis.

This system has two zeros $\pm\sqrt{\frac{T_2}{C_b I_m T_1}}$ on the real axis of the s -plane. The positive zero makes this system a non-minimum phase system. Suppose that the purpose is to redesign this system so that it becomes a minimum phase system, the system structure must be modified. It is not difficult to find that if the mechanism is reconfigured so that the transformer parameters T_1 or T_2 becomes negative, there will not be no positive zero. However, the zeros $\pm\sqrt{|\frac{T_2}{C_b I_m T_1}|}j$ will then be on the imaginary axis. The system is still a "marginally" non-minimum phase system, which will cause difficulty in optimal controller design [10]. Thus, in addition to this change, some damping resistance must be added to the zero dynamics so that the zeros can be moved to the left half plane. From the zero dynamics model in Figure 3.13, we can see that the resistance elements (R_m , R_i , R_f) are either not causally related to C_b, I_m or not involved because ω_{arm} is required to be zero. As a result, they do not play any role in the zero dynamics. So it will not help to add resistance elements at any of these locations. Also, since the junction loop is formed due to the kinematic relations, there is no physically possible way to add resistance on the bonds between ω_{arm} and ω_{inertial} , or ω_{relative} and ω_{inertial} . A possible solution is to add the resistance element parallel or serial to the capacitance element (belt flexibility). The corresponding bond graph models are shown in Figure 3.14 and Figure 3.15. By applying the ZDIP procedure,

it can be found that both designs lead to minimum phase systems. However, design 1 requires the resistance element to be built serially in the belt. In this configuration, there is no constraint to stop the belt from elongating. This is not a desired behavior. On the other hand, design 2 can be implemented by adding a parallel damping device between the motor shaft and the pinion shaft. Thus, it is a possible configuration with minimum phase zeros. The added resistance will provide the force to eliminate the non-minimum phase behavior.

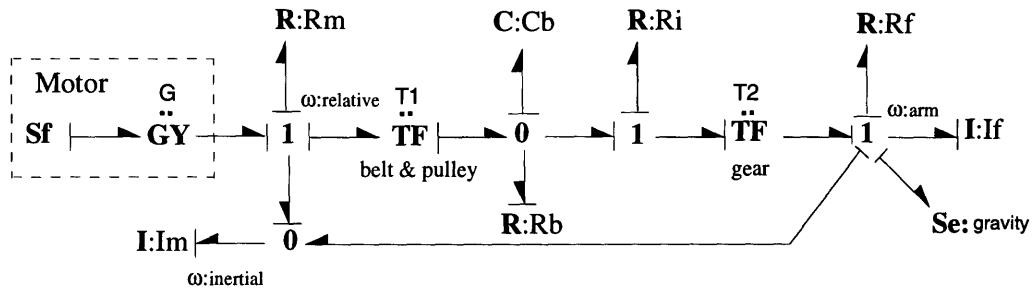


Figure 3.14: Design 1 of an arm prosthesis.

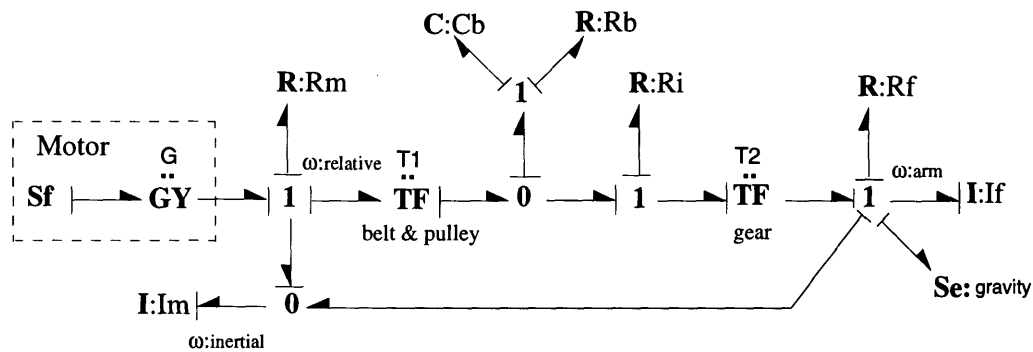


Figure 3.15: Design 2 of an arm prosthesis.

Note that if the zeros were found by calculating the transfer function, there will be no clues on how to modify the system structures. This procedure makes it possible to shift the zeros by changing system configurations without a trial-and-error process.

3.5 Conclusion

In this chapter, a Zero Dynamics Identification Procedure (ZDIP) is proposed to derive the zero dynamics of physical systems from the bond graph models. This method incorporates the derivation of zero dynamics in the differential geometric approach and the causality manipulation in the representation of bond graph models. Using the proposed procedure, the system elements which are responsible for the zero dynamics can be identified. Isolated subsystems which exhibit the zero dynamics can be found if they exist. Because the procedure utilizes the causality manipulation to partition the original systems, specific heuristic rules for the systems with specific patterns are not necessary. This approach is generalized for MIMO systems in the next chapter.

Zero Dynamics of Physical Systems from Bond Graph Models - Part II : MIMO Systems

4.1 Introduction

Zero dynamics is an important feature in system analysis and controller design. Its behavior plays a major role in determining the performance limits of certain feedback systems [10, 21, 39]. Since the intrinsic zero dynamics can not be influenced by feedback compensation, it is important to design physical systems so that they possess desired zero dynamics. However, the derivation of the zero dynamics is usually complicated even for SISO systems, especially if a form which is closely related to the physical system and suitable for design is required. Thus a method of designing the zero dynamics through structural analysis is proposed in this thesis.

The zero dynamics has an interpretation as the resultant internal dynamics when suitable initial conditions and control inputs are applied to maintain the outputs zero for all time. For simple SISO physical systems, part of or full zero dynamics may be obtained by direct inspection of the system structure [31, 47]. Although these results do not provide a systematic way for dealing with general systems, the approach of using the models at a more abstract level such as bond graph models does point out a potential direction. In the previous chapter, the definition of the zero dynamics in the differential geometric approach and the causality manipulation in the bond graph representation are incorporated. By doing so, the design of physical systems,

including the consideration of the zero dynamics, become straightforward. Since this approach does not depend on the heuristic rules for systems with specific patterns, the results can be applied for general SISO systems. In this chapter, this approach is generalized for MIMO systems.

In section 2, the definition of the zero dynamics for MIMO systems from the differential geometric approach point of view is reviewed. Section 3 discusses the vector relative degrees of MIMO systems and a dynamic extension procedure in the bond graph representation. Section 4 shows the extension of the proposed procedure for MIMO systems with a relative degree. Section 5 considers the classes of MIMO systems without a vector relative degree and the application of the mentioned dynamic extension procedure. The conclusion is given in Section 6.

4.2 Zero Dynamics in the Differential Geometric Approach

In this section, the definition of the zero dynamics for MIMO systems from the differential geometric point of view is reviewed [21, 39]. This approach deals with a class of nonlinear systems with the general form

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})\mathbf{u} \quad (4.1)$$

$$\mathbf{y} = \mathbf{h}(\mathbf{x}) \quad (4.2)$$

, where $\mathbf{u} \in R^p$ is the input vector, $\mathbf{y} \in R^m$ is the output vector and $\mathbf{x} \in R^n$ represents the state vector. The zero dynamics is defined as the internal dynamics of the system when the necessary initial conditions and control inputs are applied to keep the outputs zero for all time. So the zero dynamics can be described as

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})\mathbf{u}^* \quad (4.3)$$

with suitable initial states. The conditions for \mathbf{u}^* to exist and how the system evolves under these conditions can be derived in a rather systematic and rigorous way by considering a local coordinate transformation problem. This analysis will in turn lead to the application of input-output linearization and certain dynamic decoupling problems. In this paper, the relevant existing results are stated since the details can be found in [21, 39].

For nonlinear MIMO systems, if the system has a vector relative degree, the definition of the zero dynamics is a straightforward extension from the one for SISO systems. Assuming that an MIMO system has the same number of input and output channels, say m , this system is said to have a vector relative degree $\{r_1, r_2, \dots, r_m\}$ in the neighborhood of a point \mathbf{x}° if

(i) $L_{g_j} L_{\mathbf{f}}^k h_i(\mathbf{x}) = 0$ for all $1 \leq i \leq m$, $1 \leq j \leq m$, for all $k \leq r_i - 1$ and for all \mathbf{x} in a neighborhood of \mathbf{x}° .

(ii) The decoupling matrix
$$\begin{bmatrix} L_{g_1} L_{\mathbf{f}}^{r_1-1} h_1(\mathbf{x}) & \dots & L_{g_m} L_{\mathbf{f}}^{r_1-1} h_1(\mathbf{x}) \\ L_{g_1} L_{\mathbf{f}}^{r_2-1} h_2(\mathbf{x}) & \dots & L_{g_m} L_{\mathbf{f}}^{r_2-1} h_2(\mathbf{x}) \\ \dots & \dots & \dots \\ L_{g_1} L_{\mathbf{f}}^{r_m-1} h_m(\mathbf{x}) & \dots & L_{g_m} L_{\mathbf{f}}^{r_m-1} h_m(\mathbf{x}) \end{bmatrix}_{\mathbf{x}^\circ}$$
 is nonsingular.

Simply speaking, r_i 's are determined by differentiating each output $y_i = h_i(\mathbf{x})$ until at least one of the inputs appears. The differentiations can be represented as

$$y_i = h_i(\mathbf{x}) \quad (4.4)$$

$$\dot{y}_i = L_{\mathbf{f}} h_i(\mathbf{x}) \quad (4.5)$$

$$\dots \quad (4.6)$$

$$y_i^{(r_i-1)} = L_{\mathbf{f}}^{(r_i-1)} h_i(\mathbf{x}) \quad (4.7)$$

$$y_i^{(r_i)} = L_{\mathbf{f}}^{(r_i)} h_i(\mathbf{x}) + \sum_{j=1}^m L_{g_j} L_{\mathbf{f}}^{r_i-1} h_i(\mathbf{x}) u_j \quad (4.8)$$

The decoupling matrix is constructed by the terms $L_{g_j} L_{\mathbf{f}}^{r_i-1} h_i(\mathbf{x})$, $1 \leq j \leq m$ in Eqn.(4.8) for each output y_i , $1 \leq i \leq m$.

According to this definition, if an MIMO system has a vector relative degree in the neighborhood of a certain point, the r'_i 's have the following property

$$r_1 + r_2 + \dots + r_m = r \leq n \quad (4.9)$$

If a system has a relative degree $r = r_1 + r_2 + \dots + r_m = n$, this system has no zero dynamics. Otherwise, the system has a zero dynamics of order $n - r$. Similar to the SISO case, the zero dynamics in such an MIMO system will evolve under the following constraints if the required control and initial states are applied,

$$Z^* = \{ \mathbf{x} \in R^n : h_i(\mathbf{x}) = \dots = L_{\mathbf{f}}^{r_i-1} h_i(\mathbf{x}) = 0 \} \quad (4.10)$$

$$\text{or equivalently, } Z^* = \{ \mathbf{x} \in R^n : y_i(\mathbf{x}) = \dot{y}_i(\mathbf{x}) \dots = y_i^{(r_i-1)}(\mathbf{x}) = 0 \} \quad (4.11)$$

, where i is corresponding to the outputs, and $1 \leq i \leq m$. These constraints can be easily derived through a transformation $\mathbf{z} = \Phi(\mathbf{x})$ which converts the state equations to the following normal form,

$$\begin{aligned} \dot{z}_1^i &= z_2^i \\ \dot{z}_2^i &= z_3^i \\ &\dots \\ \dot{z}_{r_i-1}^i &= z_{r_i}^i \end{aligned} \quad (4.12)$$

$$\begin{aligned} \dot{z}_{r_i} &= a(\xi_1, \xi_2, \dots, \xi_n, \eta) + \sum_{j=1}^m b_j(\xi_1, \xi_2, \dots, \xi_n, \eta) u_j \\ \dot{\eta} &= q(\xi_1, \xi_2, \dots, \xi_n, \eta) \\ y_i &= z_1^i \end{aligned} \quad (4.13)$$

, where $1 \leq i \leq m$, and

$$\xi_1 = \begin{bmatrix} z_1^1 \\ \dots \\ z_{r_1}^1 \end{bmatrix}, \quad \xi_2 = \begin{bmatrix} z_1^2 \\ \dots \\ z_{r_2}^2 \end{bmatrix}, \quad \dots, \quad \xi_m = \begin{bmatrix} z_1^m \\ \dots \\ z_{r_m}^m \end{bmatrix}, \quad \eta = \begin{bmatrix} z_{r+1} \\ \dots \\ z_n \end{bmatrix} \quad (4.14)$$

If the initial conditions and control inputs are chosen so that the system states are initiated in the set $\xi_1 = \xi_2 = \dots = \xi_m = 0$, the system will always evolve in this set. Therefore, the zero dynamics is given by the following differential equations.

$$\dot{\eta} = q(0, 0, \dots, 0, \eta) \quad (4.15)$$

In the original coordinates, this leads to the constraints of Eqn.(4.10) or (4.11). In this paper, Eqn.(4.1) and Eqn.(4.10) or (4.11) will be used to evaluate the properties of the zero dynamics.

For MIMO systems which do not have a vector relative degree, the zero dynamics may be found by a more general zero dynamics algorithm under a weaker hypothesis [21]. However, a simpler approach is to add integrators at certain input channels so that the relative degree may be defined in terms of the nominally defined inputs. This approach is called dynamic extension because the feedback signals through these channels are not static anymore. Since the zero dynamics is left unchanged under the dynamic extension [21], the definition of the zero dynamics through a vector relative degree can still be applied for the revised model which contains the extra states from the dynamic extension. If the vector relative degree of a system can not be defined even through this approach, it can be proved that the relative degree of such a system can not be defined through any other dynamic extension of the following general form,

$$\mathbf{u} = \alpha(\zeta, x) + \beta(\zeta, x)\mathbf{v} \quad (4.16)$$

$$\dot{\zeta} = \gamma(\zeta, x) + \delta(\zeta, x)\mathbf{v} \quad (4.17)$$

In this case, the existence of a reasonably defined zero dynamics is questionable. An example of this class of systems is discussed in section 5.2.

4.3 Zero Dynamics of MIMO Systems from Bond Graph Models

In this section, the ZDIP procedure [20] is extended for MIMO systems. If an MIMO nonlinear system has a vector relative degree as defined in section 2, a similar procedure of deriving the zero dynamics can be developed by a straightforward extension of the one for SISO systems. However, for MIMO systems, the relative degree and zero dynamics are not always well-defined in terms of system configuration. Thus, the existence of relative degree and zero dynamics in MIMO systems is first discussed. The application of ZDIP procedure to MIMO systems will be shown in the following sections.

4.3.1 Vector relative degree

As described in section 4.2, the notion of relative degree can be extended to MIMO systems under the assumption that the decoupling matrix is nonsingular. The singularity of the decoupling matrix depends on the operating points, the system parameters and the system structures. In this thesis, the concern is the structural configurations which cause the relative degree of physical systems unable to be defined. It will be shown that a certain category of such systems can be modified by a dynamic extension method so that the relative degree can be rendered. The zero dynamics can then still be found by using the revised models.

The decoupling matrix is *structurally* singular if and only if it contains (1) zero columns (2) zero rows (3) dependent rows or columns, at any operating point with any set of system parameters. Note that condition 2 never happens according to the definition of the decoupling matrix itself unless part of the system is isolated and certain output is not causally related to any input. Also, if two inputs are deliberately

chosen to be at the same location, the corresponding columns will be exactly the same. This is a trivial example with condition 3. Since these cases rarely happen and can be detected easily, they are not considered in this paper. Condition 1 and the non-trivial cases of condition 3, however, are not obvious from system configurations. The following propositions describe the corresponding structures in the bond graph representation.

Proposition 4.1: *In an MIMO bond graph model, if any one of the inputs can not be connected to at least one output with a shortest causal path, under the condition that there are no other inputs that can be connected to this output with a "shorter" causal path, in other words, if there are always other inputs "closer"¹ to all the outputs, then this system does not have a vector relative degree.*

Proposition 4.2: *In an MIMO bond graph model, if any two of the shortest causal paths between the input and output pairs share any of the junctions, dissipative elements, or energy storage elements, (i.e. these two causal paths partially overlap) and there is no other alternative causality assignments, choice of shortest causal paths to avoid this overlapping, then this system does not have a vector relative degree.*

Proofs for both propositions are given in Appendix B.3

4.3.2 Dynamic extension procedure

The system configurations described in proposition 4.1 cause the decoupling matrix to be singular because of condition 1 (zero columns). This condition can be eliminated if integrators are added at certain input channels to increase the order of the dynamics between these inputs and the outputs[21]. As mentioned in section 2, this procedure is one form of dynamic extension. Using the bond graph interpretation of the relative degree for each input-output pair[20], a corresponding bond graph procedure for

¹"Shorter" or "closer" is defined in the sense that the causal path contains a smaller number of energy storage elements with integral causality.

dynamic extension is derived and stated as follows,

Dynamic extension procedure:

- (1) *Determine the shortest causal path between each output, $1 \leq i \leq m$, to the "nearest" input² and denote the number of energy storage elements with integral causality on these shortest causal paths as n'_i s. Denote the unconnected inputs³ as u'_j s.*
- (2) *Determine the shortest causal path between each unconnected input u_j to the nearest outputs i_{uj} and denote the number of integral causality on the shortest causal paths as n'_{ij} s.*
- (3) *Add $n_{ij} - n_i$ integrators at the control channels which are originally connected to output i'_{uj} s by the shortest causal paths in step 1. The inputs to the integrators now are treated as the new system inputs.*
- (4) *Reapply step 1 for the modified model. In this process, there should be at least two shortest causal paths for each output i'_{uj} s. One leads to the control channel in step 3 where integrators are added. Another one leads to an originally unconnected inputs u_j in step 2. The latter path should be selected as the shortest causal path for output i_{uj} . By doing so, the unconnected inputs u'_j s are paired with i'_{uj} s and the system configuration in proposition 4.1 does not exist.*

Remark : After applying the dynamic extension procedure, the structural condition in proposition 4.2 may still exist. Due to condition 3 (dependent rows), the decoupling matrix of such a system is always singular. In this case, the relative degree can not be defined through any dynamic extension. From design point of view, this is an undesirable structure with inherent difficulty in controller implementation. Thus, as a guideline, the inputs and the outputs of MIMO systems should avoid such a

²"Nearest" is defined in the sense that the causal path contains the least number of integral causality.

³The inputs which are not connected by any shortest causal path after this process are called unconnected inputs.

configuration.

In this chapter, the systems or the revised models which have vector relative degrees in terms of system configuration are considered as the candidates for the ZDIP procedure.

4.4 MIMO Systems with a Vector Relative Degree

If an MIMO nonlinear system has a vector relative degree as defined in section 2, the relative degree r_i associated with output y_i is indicated by the number of integral causality on the shortest causal path which connects this output to its "nearest" input. This is a direct extension of the rules for SISO systems [20]. Using these shortest causal paths, the zero dynamics can be identified by a procedure similar to the SISO ZDIP procedure. A revised version of the ZDIP procedure for MIMO systems is listed as follows.

ZDIP procedure for MIMO systems

- (1) *Apply the normal Sequential Causality Assignment Procedure (SCAP) to the system bond graph model.*
- (2) *Determine the non-overlapping causal paths which define the vector relative degree $\{r_1, r_2, \dots, r_m\}$ ⁴.*
- (3) *Remove the bond which is related to the output variables, and assign a zero value to the junctions which these bonds are connected to. This indicates that the common efforts of the 0 junctions or the common flows of the 1 junctions are zero.*
- (4) *Reverse the causality on the bonds which are connected to the inputs.*
- (5) *Replace every integral causality on these shortest causal paths by a derivative*

⁴The order of the zero dynamics can be determined by the number of energy storage elements with integral causality not on any of these paths.

causality.

- (6) *Keep every elements with integral causality not on these shortest causal path unchanged. These are the candidate elements which will contribute the independent state variables for the zero dynamics.*
- (7) *Complete the causality assignment without violating the junction constraints.*
- (8) *Derive the zero dynamics according to the causality assignment.*
- (9) *If the above derivation requires the causal outputs from the elements whose causality has been reversed, solve for these variables using the algebraic equations in Eqn.(4.10) so that they can be represented by the independent state variables determined in step 6.*

Example : Consider the MIMO system of Figure 4.1. The inputs and the outputs of this system are u_1, u_2 and f_1, f_2 respectively. By searching the causal paths, since u_1 is much "closer" to f_1 than u_2 and u_2 is much "closer" to f_2 than u_1 , the decoupling matrix is a diagonal matrix. It is clear that this system has a vector relative degree in terms of system configuration. So the procedure can be executed as shown in Figure 4.2. The resultant model shows that the zero dynamics of this system can be partitioned as a first order system in subsystem 1 and a 3rd order system in subsystem 2 as shown in Figure 4.3. The dashed line with an arrow indicates that the input variable (effort) at port b_2 of subsystem 2 is determined by the output variable at port b_1 of the same subsystem. The other input variables at port b_1, b_3 and b_4 are all zero. Thus these two subsystems are isolated. Their properties can be studied separately no matter whether the system is linear or nonlinear.

Assuming that the constitutive equations for C_i, I_i and R_i are $e_{C_i} = \psi_{C_i}(q_{C_i})$, $f_{I_i} = \psi_{I_i}(p_{I_i})$, $e_{R_i} = \psi_{R_i}(f_{R_i})$ respectively, the state equations of the zero dynamics

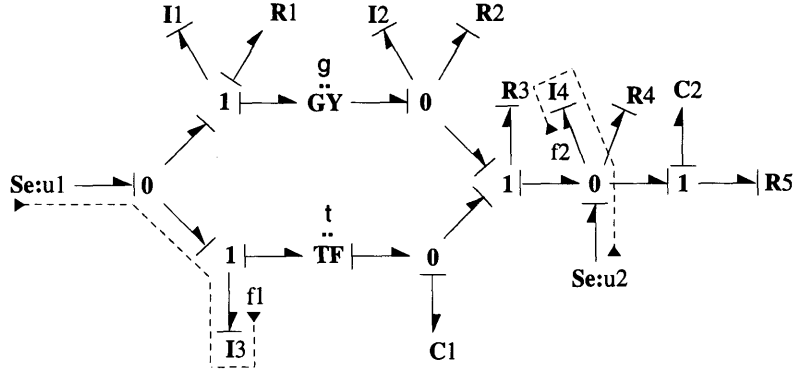


Figure 4.1: Bond graph model of an MIMO system.

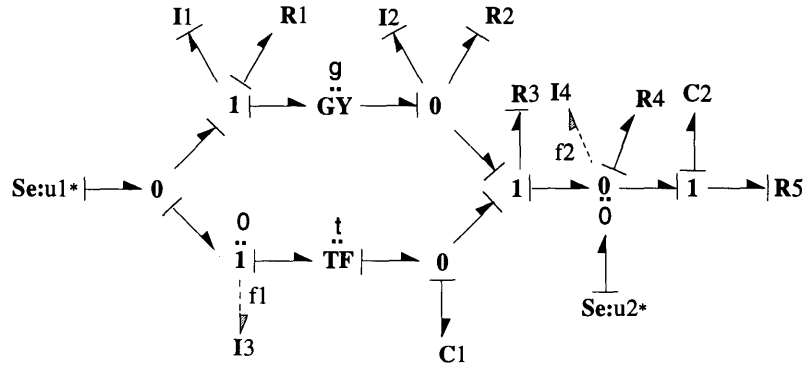


Figure 4.2: The causality assignment after ZDIP procedure.

can be derived as follows

$$\dot{q}_{C2} = \psi_{R5}^{-1}(-\psi_{C2}(q_{C2}))$$

$$\dot{q}_{C1} = -\psi_{R3}^{-1}(\psi_{C1}(q_{C1}) + \psi_{I1}(p_{I1})g)$$

$$\begin{aligned} \dot{p}_{I1} = & \psi_{C1}(q_{C1})t - \psi_{R1}(\psi_{I1}(p_{I1})) - g\psi_{I2}(p_{I2}) \\ & - g \left[\psi_{R3}^{-1}(\psi_{C1}(q_{C1}) + \psi_{I1}(p_{I1})g) + \psi_{R2}^{-1}(\psi_{I1}(p_{I1})g) \right] \end{aligned}$$

$$\dot{p}_{I2} = \psi_{I1}(p_{I1})g$$

With linear constitutive equations $e_{C_i} = \frac{q_{C_i}}{C_i}$, $f_{I_i} = \frac{p_{I_i}}{I_i}$, $e_{R_i} = f_{R_i}R_i$, the state equations can be written as

$$\dot{q}_{C2} = -\frac{q_{C2}}{R_5 C_2}$$

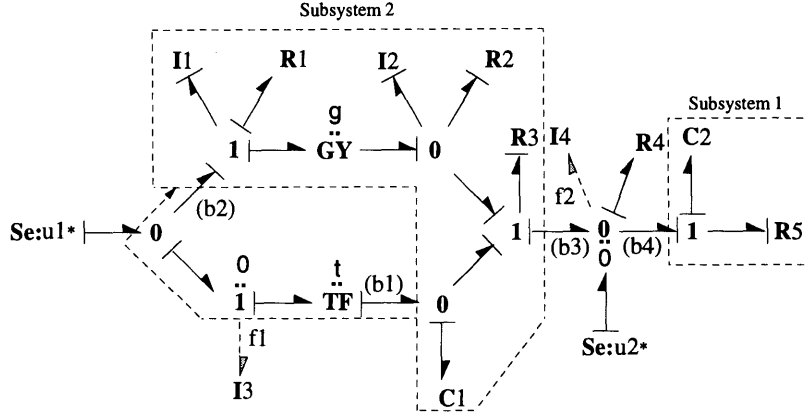


Figure 4.3: The zero dynamics identified by ZDIP procedure.

$$\begin{bmatrix} \dot{q}_{C1} \\ \dot{p}_{I1} \\ \dot{p}_{I2} \end{bmatrix} = \begin{bmatrix} \frac{-1}{R_3 C_1} & & & & 0 \\ \frac{t}{C_1} - \frac{1}{R_3 C_1} & -\frac{R_1}{I_1} - \frac{g^2}{R_3 I_1} - \frac{g^2}{R_2 I_1} & & & \frac{-g}{I_2} \\ 0 & \frac{g}{I_1} & & & 0 \end{bmatrix} \begin{bmatrix} q_{C1} \\ p_{I1} \\ p_{I2} \end{bmatrix}$$

Thus, $\frac{-1}{R_3 C_1}$ and the eigenvalues of the matrix in the above equation would be the transmission zeros of the overall MIMO system.

Note that if the equations are transformed into the normal form of Eqn.(4.12), the representation of the isolated subsystems is not revealed. The properties (e.g. stability) of such a high order nonlinear zero dynamics would be difficult to determine, not to mention the effort of finding the transformation functions. Even for linear systems, the calculation of the MIMO transmission zeros also requires significant efforts. By this procedure, the zero dynamics of MIMO systems can be derived in a systematic way. The zeros can be found by solving the eigenvalues of the lower-order subsystems. This approach can reduce the effort of computation and the risk of unstable numerical errors in addition to its design purpose.

4.5 MIMO Systems without a Vector Relative Degree

If an MIMO nonlinear system does not have a vector relative degree, the dynamic extension procedure in section 4.3.2 may be used to revise the model. If the revised model has a vector relative degree in terms of system configuration, the ZDIP procedure can be applied in the same manner as shown in the previous section. The following examples illustrate the use of the dynamic extension procedure. To explore the relations between the proposed approach and the traditional methods, three approaches will be used to derive the system's zeros (zero dynamics) for each example : numerical computations, differential geometric derivation and the proposed ZDIP procedure using the bond graph representation.

4.5.1 Systems with a relative degree under dynamic extension

In this subsection, a typical example of the class of systems with a relative degree under dynamic extension is discussed. Consider a simple mass-damper-spring system shown in Figure 4.4. The inputs are $u_1 = F_1$, $u_2 = F_2$ and the outputs are $y_1 = \dot{x}_1$, $y_2 = \dot{x}_2$.

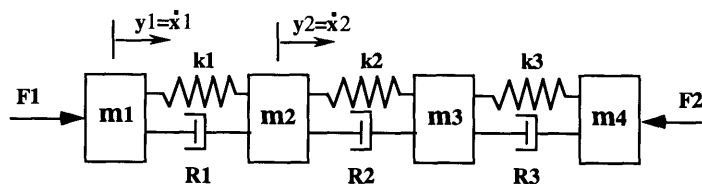


Figure 4.4: A 2-input 2-output mechanical system.

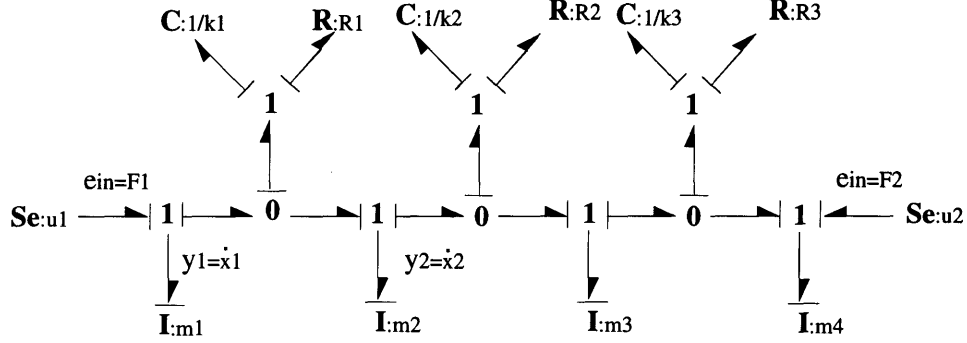


Figure 4.5: A corresponding bond graph model.

Numerical computation

For simplicity, all elements are assumed linear⁵. The velocities associated with the masses m_1 , m_2 , m_3 and m_4 are f_1 , f_2 , f_3 and f_4 respectively. The forces associated with the springs C_1 , C_2 and C_3 are e_1 , e_2 and e_3 . According to the bond graph model in Figure 4.5, the state equations are derived as follows.

$$\dot{f}_4 = \frac{1}{m_4}[e_3 - (f_3 - f_4)R_3 - u_2] \quad (4.18)$$

$$\dot{e}_3 = (f_3 - f_4)k_3 \quad (4.19)$$

$$\dot{f}_3 = \frac{1}{m_3}[e_2 + (f_2 - f_3)R_2 - e_3 - (f_3 - f_4)R_3] \quad (4.20)$$

$$\dot{e}_2 = (f_2 - f_3)k_2 \quad (4.21)$$

$$\dot{f}_2 = \frac{1}{m_2}[e_1 + (f_1 - f_2)R_1 - e_2 - (f_2 - f_3)R_2] \quad (4.22)$$

$$\dot{e}_1 = (f_1 - f_2)k_1 \quad (4.23)$$

$$\dot{f}_1 = \frac{1}{m_1}[u_1 - e_1 - (f_1 - f_2)R_1] \quad (4.24)$$

$$y_1 = f_1 \quad (4.25)$$

⁵Using the ZDIP procedure, the system does not have to be linear.

$$y_2 = f_2 \quad (4.26)$$

In a matrix form, the state equations can be written as follows.

$$\dot{\mathbf{x}} = \mathbf{Ax} + \mathbf{Bu}$$

$$\mathbf{y} = \mathbf{Cx} + \mathbf{Du}$$

$$\Rightarrow \begin{pmatrix} \dot{f}_4 \\ \dot{e}_3 \\ \dot{f}_3 \\ \dot{e}_2 \\ \dot{f}_2 \\ \dot{e}_1 \\ \dot{f}_1 \end{pmatrix} = \begin{bmatrix} \frac{-R_3}{m_4} & \frac{1}{m_4} & \frac{R_3}{m_4} & 0 & 0 & 0 & 0 \\ -k_3 & 0 & k_3 & 0 & 0 & 0 & 0 \\ \frac{R_3}{m_3} & \frac{-1}{m_3} & \frac{-R_2-R_3}{m_3} & \frac{1}{m_3} & \frac{R_2}{m_3} & 0 & 0 \\ 0 & 0 & -k_2 & 0 & k_2 & 0 & 0 \\ 0 & 0 & \frac{R_2}{m_2} & \frac{-1}{m_2} & \frac{-R_1-R_2}{m_2} & \frac{1}{m_2} & \frac{R_1}{m_2} \\ 0 & 0 & 0 & 0 & -K_1 & 0 & k_1 \\ 0 & 0 & 0 & 0 & \frac{R_1}{m_1} & \frac{-1}{m_1} & \frac{-R_1}{m_1} \end{bmatrix} \begin{pmatrix} f_4 \\ e_3 \\ f_3 \\ e_2 \\ f_2 \\ e_1 \\ f_1 \end{pmatrix} + \begin{bmatrix} 0 & \frac{-1}{m_4} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ \frac{1}{m_1} & 0 \end{bmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{pmatrix} f_4 \\ e_3 \\ f_3 \\ e_2 \\ f_2 \\ e_1 \\ f_1 \end{pmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

For the convenience of calculation, the following numerical values are selected. $R_1 = 1 \frac{N \cdot sec}{m}$, $R_2 = 2 \frac{N \cdot sec}{m}$, $R_3 = 4 \frac{N \cdot sec}{m}$, $k_1 = 4 \frac{N}{m}$, $k_2 = 2 \frac{N}{m}$, $k_3 = 8 \frac{N}{m}$, $m_1 = m_2 = m_3 = m_4 = 1Kg$.

One way to obtain the system's transmission zeros is by calculating the determinant of the system's transfer function matrix. For this system, the transfer function matrix is calculated as follows.

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \mathbf{G}(s) \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

$$\mathbf{G}(s) = \frac{1}{D(s)} \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix}$$

$$= \frac{1}{D(s)} \begin{bmatrix} s^6 + 13s^5 + 58s^4 + 138s^3 + 176s^2 + 112s + 64 & -8s^3 - 56s^2 - 112s - 64 \\ s^5 + 14s^4 + 66s^3 + 128s^2 + 112s + 64 & -8s^4 - 32s^3 - 72s^2 - 112s - 64 \end{bmatrix}$$

,where $D(s) = s^7 + 14s^6 + 74s^5 + 230s^4 + 424s^3 + 448s^2 + 256s$. The zeros of the determinant of $\mathbf{G}(s)$ are $-7.1429, -2.1888, -1.5381 \pm 2.3328i, -0.7959 \pm 1.2097i, -2, -1, 0, 0$. Note that $-7.1429, -2.1888, -1.5381 \pm 2.3328i, -0.7959 \pm 1.2097i$ and 0 are just the repeats of the poles. Therefore, the transmission zeros of this system are $0, -1$ and -2 . This result can also be obtained by solving the generalized eigenvalue problem,

$$\begin{bmatrix} s\mathbf{I} - \mathbf{A} & -\mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{pmatrix} \mathbf{x}_0 \\ - \\ - \\ \mathbf{u}_0 \end{pmatrix} = \mathbf{0} \quad (4.27)$$

,where the $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ matrices are defined as above, \mathbf{x}_0 is the corresponding initial condition vector and \mathbf{u}_0 is the corresponding input amplitude vector. The eigenvectors of the generalized eigenvalues $0, -1$ and -2 are listed as follows.

$$\begin{pmatrix} 0 \\ 0.4472 \\ 0 \\ 0.4472 \\ 0 \\ 0.4472 \\ 0 \\ - \\ 0.4472 \\ 0.4472 \end{pmatrix}, \begin{pmatrix} 0.3162 \\ 0.5060 \\ 0.2530 \\ 0.5060 \\ 0 \\ 0 \\ 0 \\ - \\ 0 \\ 0.5692 \end{pmatrix}, \begin{pmatrix} 0.2182 \\ 0.8729 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ - \\ 0 \\ 0.4364 \end{pmatrix}$$

Note that these eigenvectors show the nontrivial eigenmodes of the system motion, where both the outputs are kept zero for all time. The first mode shows that when $e_{10} = 0.4472, e_{20} = 0.4472, e_{30} = 0.4472$ and all the initial velocities of the masses are zero, if $u_1 = 0.4472$ and $u_2 = 0.4472$ are applied, the outputs will be kept zero for all time⁶. The corresponding generalized eigenvalue of this mode is 0 . This is obviously the case when the system achieves an static force equilibrium and all the

⁶The absolute amplitude of these numbers are of course not important, if only the ratios are kept the same.

masses have a zero velocity for all time. Similarly, the second mode shows that when $f_{40} = 0.3162$, $f_{30} = 0.2530$, $e_{30} = 0.5060$, $f_{20} = 0.5060$ and other states are zero, if $u_1 = 0$ and $u_2 = 0.5692e^{-1}$ are applied, the outputs will be kept zero for all time. The third mode shows that when $f_{40} = 0.2182$, $e_{30} = 0.8729$, and other states are zero, if $u_1 = 0$ and $u_2 = 0.4364e^{-2}$ are applied, the outputs will be kept zero for all time.

Differential geometric approach:

According to the differential geometric approach, the zero dynamics will evolve on the set defined by Eqn.(4.10) or (4.11) if the system has a vector relative degree. In other words, if the constraints in Eqn.(4.10) or (4.11) are imposed on the system dynamics, the residual dynamics will be the zero dynamics. So first, the system's relative degrees are examined by taking successive differentiation of the outputs.

$$y_1 = f_1 \quad (4.28)$$

$$\dot{y}_1 = \dot{f}_1 = \frac{1}{m_1}[u_1 - e_1 - (f_1 - f_2)R_1] \quad (4.29)$$

$$y_2 = f_2 \quad (4.30)$$

$$\dot{y}_2 = \dot{f}_2 = \frac{1}{m_2}[e_1 + (f_1 - f_2)R_1 - e_2 - (f_2 - f_3)R_2] \quad (4.31)$$

$$\ddot{y}_2 = \ddot{f}_2 = \frac{1}{m_2}[\dot{e}_1 + (\dot{f}_1 - \dot{f}_2)R_1 - \dot{e}_2 - (\dot{f}_2 - \dot{f}_3)R_2] \quad (4.32)$$

Note that \dot{f}_1 , which is a function of u_1 , appears in Eqn.(4.32). Therefore, the decoupling matrix is

$$\begin{bmatrix} \frac{1}{m_1} & 0 \\ \frac{R_1}{m_1 m_2} & 0 \end{bmatrix} \quad (4.33)$$

The zero column shows that the input u_2 does not appear in the differentiations of the outputs. One way to cope with this problem is using the dynamic extension procedure. For this example, an integrator added at input u_1 will suffice.

$$\zeta = u_1 \quad (4.34)$$

$$u_1^* = \dot{\zeta} \quad (4.35)$$

The nominal input u_1^* is treated as an input and u_1 is treated as a state. For this new system, the differentiations of the outputs are

$$y_1 = f_1 \quad (4.36)$$

$$\dot{y}_1 = \dot{f}_1 = \frac{1}{m_1}[\dot{\zeta} - \dot{e}_1 - (f_1 - f_2)R_1] \quad (4.37)$$

$$\begin{aligned} \ddot{y}_1 &= \ddot{f}_1 = \frac{1}{m_1}[\ddot{\zeta} - \ddot{e}_1 - (\dot{f}_1 - \dot{f}_2)R_1] \\ &= \frac{1}{m_1}[u_1^* - \dot{e}_1 - (f_1 - f_2)R_1] \end{aligned} \quad (4.38)$$

and

$$y_2 = f_2 \quad (4.39)$$

$$\dot{y}_2 = \dot{f}_2 = \frac{1}{m_2}[e_1 + (f_1 - f_2)R_1 - e_2 - (f_2 - f_3)R_2] \quad (4.40)$$

$$\ddot{y}_2 = \ddot{f}_2 = \frac{1}{m_2}[\dot{e}_1 + (\dot{f}_1 - \dot{f}_2)R_1 - \dot{e}_2 - (\dot{f}_2 - \dot{f}_3)R_2] \quad (4.41)$$

$$y_2^{(3)} = f_2^{(3)} = \frac{1}{m_2}[\ddot{e}_1 + (\ddot{f}_1 - \ddot{f}_2)R_1 - \ddot{e}_2 - (\ddot{f}_2 - \ddot{f}_3)R_2] \quad (4.42)$$

Note that in Eqn.(4.42), \ddot{f}_1 contains u_1^* ; \ddot{f}_3 contains \dot{f}_4 , which is a function of u_2 .

$$\ddot{f}_1 = \frac{1}{m_1}[u_1^* - \dot{e}_1 - (f_1 - f_2)R_1] \quad (4.43)$$

$$\ddot{f}_3 = \frac{1}{m_3}[\dot{e}_2 + (\dot{f}_2 - \dot{f}_3)R_2 - \dot{e}_3 - (f_3 - f_4)R_3] \quad (4.44)$$

Therefore, the new decoupling matrix is now

$$\begin{bmatrix} \frac{1}{m_1} & 0 \\ \frac{R_1}{m_1 m_2} & \frac{-R_2 R_3}{m_2 m_3 m_4} \end{bmatrix} \quad (4.45)$$

By this derivation, r_1 is found as 2 from Eqn. (4.38) and r_2 is found as 3 Eqn. (4.42).

Since $n^* - r = n^* - (r_1 + r_2) = 8 - 5 = 3$, the zero dynamics of this system should be order of 3. By substituting $y_1 = \dot{y}_1 = 0$ and $y_2 = \dot{y}_2 = \ddot{y}_2 = 0$ into the state

equations, the residual dynamics will have eigenvalues $0, -1, -2$. This is consistent with the numerical computation shown above.

Remark: By comparing these two methods, it is found that the numerical calculation considers each transmission zero separately with a corresponding eigenmode motion, while the differential geometric approach considers the zero dynamics as a whole (a subset of the system dynamics). Therefore, in the differential geometric approach, a dynamic extension at u_1 is necessary. This means that the information of \dot{u}_1 is necessary to compute the required control inputs so both the outputs can be kept zero for all time.

ZDIP procedure

The proposed ZDIP procedure is based on the differential geometric approach. However, the complexity of manipulation is drastically reduced by the bond graph representation. For this given system, the shortest causal paths from the outputs to the "nearest" inputs are shown in Figure 4.6. Since input u_2 is not connected to any output, the relative degree can not be defined. This can be fixed by the dynamic extension procedure as shown in Figure 4.7, where an integrator has been added before u_1 . The relative degree can be defined now as $r_1 = 2, r_2 = 3$ by the causal paths from \hat{u}_1 to y_1 and u_2 to y_2 . Including the integrator at channel u_1 , the order of this system is 8. Therefore the order of the zero dynamics must be 3. Using the ZDIP procedure, the subsystems which represent the zero dynamics are shown in Figure 4.8. As partitioned in the figure, subsystems 1, 2 and 3 are in a chain form⁷. The input flow to subsystem 1 is zero because y_1 and y_2 are kept zero. So this subsystem is isolated. The input effort to subsystem 2 is from the output of subsystem 1. The

⁷Using the constraints in Eqn.(4.10), the causal output from I_3 can be represented by the states in subsystem 1 and 2.

input effort to subsystem 3 is from subsystem 2 and the causal output from I_3 . For linear systems, the one way interactions between subsystems do not influence the eigenvalues. Therefore, the transmission zeros of this system are 0 , $\frac{-k_2}{R_2}$ and $\frac{-k_3}{R_3}$.

Remark 1: By inspecting the decoupling matrix, it is found that the shortest causal paths between the inputs and the outputs have a strong relation with the components in the matrix. In this example, the parameters of the physical elements which are on the shortest causal path appear in the corresponding components of the decoupling matrix. This indicates that the proposed approach effectively explores the structural information of physical systems with a systematic procedure.

Remark 2: Using the bond graph representation, alternative approaches may be found to ensure the existence of the system relative degree. For example, instead of adding an integrator at u_1 , one might increase the order of dynamics between u_1 and y_1 by considering the actuator dynamics at channel u_1 . From the differential geometric point of view, this is just another type of dynamic extension. However, from design point of view, one can modify the model or system design to facilitate dynamic analysis or controller design. This is usually not possible by using only equations.

4.5.2 Systems with ill-designed input-output configurations

In this subsection, an example of the class of systems with the structural condition of Proposition 4.2 is discussed. This example reveals the physical meaning of the described ill-designed configurations. Consider a simple mass-damper-spring system shown in Figure 4.9. The inputs are now $u_1 = F_1$, $u_2 = F_2$ and the outputs are $y_1 = \dot{x}_2$, $y_2 = \dot{x}_3$.

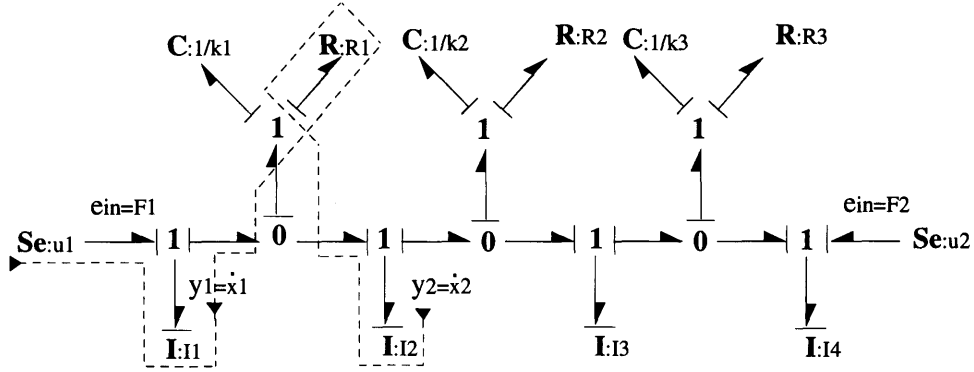


Figure 4.6: The shortest causal paths of an MIMO model.

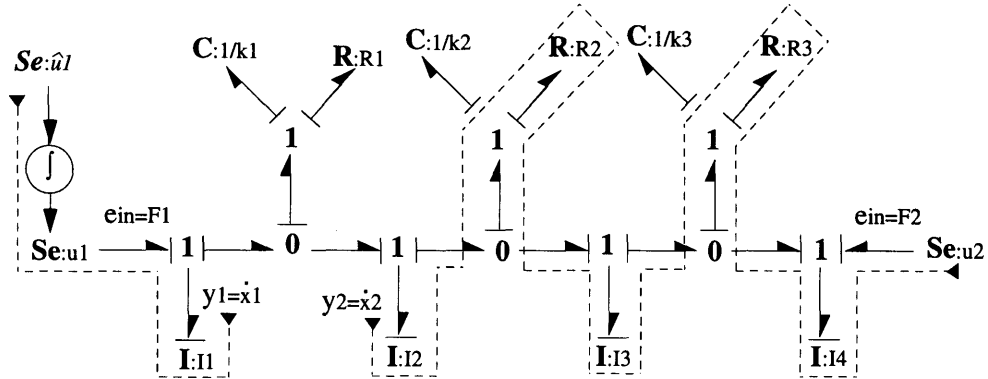


Figure 4.7: The shortest causal paths of an MIMO model by dynamic extension.

Numerical computation

For simplicity, all elements are assumed linear. The velocities associated with masses m_1 , m_2 , m_3 and m_4 are f_1 , f_2 , f_3 and f_4 respectively. The forces associated with the springs C_1 , C_2 and C_3 are e_1 , e_2 and e_3 . According to the bond graph model in Figure 4.10, the state equations are derived and found to be,

$$\dot{f}_4 = \frac{1}{m_4} [e_3 - (f_3 - f_4)R_3] \quad (4.46)$$

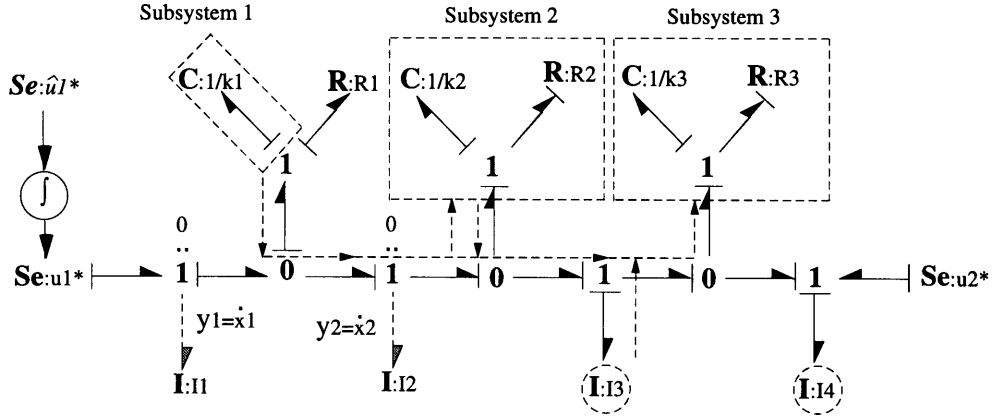


Figure 4.8: The zero dynamics identified by ZDIP procedure.

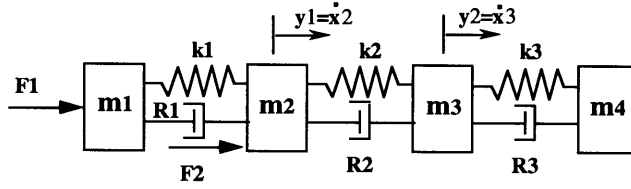


Figure 4.9: A 2-input 2-output mechanical system.

$$\dot{e}_3 = (f_3 - f_4)k_3 \quad (4.47)$$

$$\dot{f}_3 = \frac{1}{m_3}[e_2 + (f_2 - f_3)R_2 - e_3 - (f_3 - f_4)R_3] \quad (4.48)$$

$$\dot{e}_2 = (f_2 - f_3)k_2 \quad (4.49)$$

$$\dot{f}_2 = \frac{1}{m_2}[e_1 + (f_1 - f_2)R_1 - e_2 - (f_2 - f_3)R_2 + u_2] \quad (4.50)$$

$$\dot{e}_1 = (f_1 - f_2)k_1 \quad (4.51)$$

$$\dot{f}_1 = \frac{1}{m_1}[u_1 - e_1 - (f_1 - f_2)R_1] \quad (4.52)$$

$$y_1 = f_2 \quad (4.53)$$

$$y_2 = f_3 \quad (4.54)$$

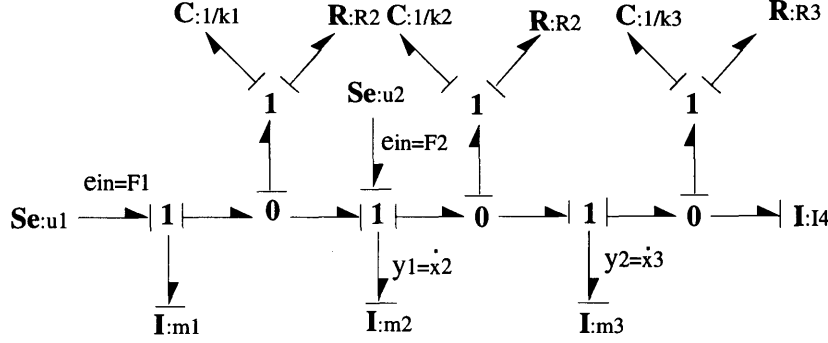


Figure 4.10: A corresponding bond graph model.

In a matrix form, the state equations can be written as follows.

$$\begin{pmatrix} \dot{f}_4 \\ \dot{e}_3 \\ \dot{f}_3 \\ \dot{e}_2 \\ \dot{f}_2 \\ \dot{e}_1 \\ \dot{f}_1 \end{pmatrix} = \begin{bmatrix} \frac{-R_3}{m_4} & \frac{1}{m_4} & \frac{R_3}{m_4} & 0 & 0 & 0 & 0 \\ -k_3 & 0 & k_3 & 0 & 0 & 0 & 0 \\ \frac{R_3}{m_3} & \frac{-1}{m_3} & \frac{-R_2-R_3}{m_3} & \frac{1}{m_3} & \frac{R_2}{m_3} & 0 & 0 \\ 0 & 0 & -k_2 & 0 & k_2 & 0 & 0 \\ 0 & 0 & \frac{R_2}{m_2} & \frac{-1}{m_2} & \frac{-R_1-R_2}{m_2} & \frac{1}{m_2} & \frac{R_1}{m_2} \\ 0 & 0 & 0 & 0 & -K_1 & 0 & k_1 \\ 0 & 0 & 0 & 0 & \frac{R_1}{m_1} & \frac{-1}{m_1} & \frac{-R_1}{m_1} \end{bmatrix} \begin{pmatrix} f_4 \\ e_3 \\ f_3 \\ e_2 \\ f_2 \\ e_1 \\ f_1 \end{pmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & \frac{1}{m_2} \\ 0 & 0 \\ \frac{1}{m_1} & 0 \end{bmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{pmatrix} f_4 \\ e_3 \\ f_3 \\ e_2 \\ f_2 \\ e_1 \\ f_1 \end{pmatrix}$$

For the convenience of calculation, the following numerical values are given. $R_1 = 1 \frac{N \cdot sec}{m}$, $R_2 = 2 \frac{N \cdot sec}{m}$, $R_3 = 4 \frac{N \cdot sec}{m}$, $k_1 = 4 \frac{N}{m}$, $k_2 = 2 \frac{N}{m}$, $k_3 = 8 \frac{N}{m}$, $m_1 = m_2 = m_3 = m_4 = 1Kg$.

One way to obtain the system's transmission zeros is by calculating the determinant of the system's transfer function matrix. For this system, the transfer function

matrix is calculated as follows.

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \mathbf{G}(s) \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

$$\begin{aligned} \mathbf{G}(s) &= \frac{1}{D(s)} \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix} \\ &= \frac{1}{D(s)} \begin{bmatrix} s^5 + 14s^4 + 66s^3 + 128s^2 + 112s + 64 & s^6 + 11s^5 + 40s^4 + 90s^3 + 144s^2 + 112s + 64 \\ 2s^4 + 18s^3 + 64s^2 + 112s + 64 & 2s^5 + 12s^4 + 42s^3 + 80s^2 + 112s + 64 \end{bmatrix} \end{aligned}$$

, where $D(s) = s^7 + 14s^6 + 74s^5 + 230s^4 + 424s^3 + 448s^2 + 256s$. However, in this example, the determinant of the transfer function matrix is identically zero: $\det[\mathbf{G}(s)] = 0$, $\forall s$. Also, it is found that the whole s space is the solution of the generalized eigenvalue problem. This indicates that the transmission zeros of this system can not be reasonably defined. However, it is not obvious from the computation procedure which part of the system structure causes this problem.

Differential geometric approach

According to the differential geometric approach, the zero dynamics will evolve on the set defined by Eqn.(4.10) or (4.11) if the system has a vector relative degree. In other words, if the constraints in Eqn.(4.10) or (4.11) are imposed on the system dynamics, the residual dynamics will be the zero dynamics. So first, the system's relative degrees are examined by taking successive differentiation of the outputs,

$$y_1 = f_2 \tag{4.55}$$

$$\dot{y}_1 = \dot{f}_2 = \frac{1}{m_2} [e_1 + (f_1 - f_2)R_1 - e_2 - (f_2 - f_3)R_2 + u_2] \tag{4.56}$$

$$y_2 = f_3 \tag{4.57}$$

$$\dot{y}_2 = \dot{f}_3 = \frac{1}{m_3} [e_2 + (f_2 - f_3)R_2 - e_3 - (f_3 - f_4)R_3] \tag{4.58}$$

$$\ddot{y}_2 = \ddot{f}_3 = \frac{1}{m_3} [\dot{e}_2 + (\dot{f}_2 - \dot{f}_3)R_2 - \dot{e}_3 - (\dot{f}_3 - \dot{f}_4)R_3] \tag{4.59}$$

Note that \dot{f}_2 , which is a function of u_2 , appears in Eqn.(4.59). Therefore, the decoupling matrix is

$$\begin{bmatrix} 0 & \frac{1}{m_2} \\ 0 & \frac{R_2}{m_2 m_3} \end{bmatrix} \quad (4.60)$$

The zero column shows that the input u_1 does not appear in the differentiations of the outputs. One way to cope with this problem is using the dynamic extension procedure. For this example, an integrator added at input u_2 will suffice to eliminate the zero column,

$$\zeta = u_2 \quad (4.61)$$

$$u_2^* = \dot{\zeta} \quad (4.62)$$

The nominal input u_2^* is considered as an input and u_2 is now treated as a state. For this new system, the differentiations of the outputs are

$$y_1 = f_2 \quad (4.63)$$

$$\dot{y}_1 = \dot{f}_2 = \frac{1}{m_2}[e_1 + (f_1 - f_2)R_1 - e_2 - (f_2 - f_3)R_2 + \zeta] \quad (4.64)$$

$$\begin{aligned} \ddot{y}_1 &= \ddot{f}_2 = \frac{1}{m_2}[\dot{e}_1 + (\dot{f}_1 - \dot{f}_2)R_1 - \dot{e}_2 - (\dot{f}_2 - \dot{f}_3)R_2 + \dot{\zeta}] \\ &= \frac{1}{m_2}[\dot{e}_1 + (\dot{f}_1 - \dot{f}_2)R_1 - \dot{e}_2 - (\dot{f}_2 - \dot{f}_3)R_2 + u_2^*] \end{aligned} \quad (4.65)$$

and

$$y_2 = f_3 \quad (4.66)$$

$$\dot{y}_2 = \dot{f}_3 = \frac{1}{m_3}[e_2 + (f_2 - f_3)R_2 - e_3 - (f_3 - f_4)R_3] \quad (4.67)$$

$$\ddot{y}_2 = \ddot{f}_3 = \frac{1}{m_3}[\dot{e}_2 + (\dot{f}_2 - \dot{f}_3)R_2 - \dot{e}_3 - (\dot{f}_3 - \dot{f}_4)R_3] \quad (4.68)$$

$$y_2^{(3)} = f_3^{(3)} = \frac{1}{m_3}[\ddot{e}_2 + (\ddot{f}_2 - \ddot{f}_3)R_2 - \ddot{e}_3 - (\ddot{f}_3 - \ddot{f}_4)R_3] \quad (4.69)$$

Note that in Eqn.(4.65), \ddot{f}_2 contains u_2^* and \dot{f}_1 , where \dot{f}_1 is a function of u_1 . Also, in Eqn.(4.69), $f_3^{(3)}$ contains \ddot{f}_2 . Therefore, although the decoupling matrix does not

have any zero column as shown below, it is still structurally singular.

$$\begin{bmatrix} \frac{R_1}{m_1 m_2} & \frac{1}{m_2} \\ \frac{R_1 R_2}{m_1 m_2 m_3} & \frac{R_2}{m_2 m_3} \end{bmatrix} \quad (4.70)$$

This is due to the fact that both outputs finally reach the inputs through the term \ddot{f}_2 . Even if this system contains nonlinear elements, the result will not change. Therefore, the zero dynamics can not be identified by this procedure.

ZDIP procedure

The proposed ZDIP procedure is based on the differential geometric approach, therefore no more results can be obtained than what has been shown above. However, using the bond graph representation, one can explore the physical meaning of the above results and predict them by simple system structural analysis. As a result, a simple solution can be found by relocating the inputs.

The shortest causal paths from the outputs to the "nearest" inputs are shown in Figure 4.11. Since input u_1 is not connected to any output through these paths, the relative degree can not be defined. This can be fixed by the dynamic extension procedure as shown in Figure 4.12. However, the causal path between u_1 and y_1 and the one between \hat{u}_2 and y_2 share the energy storage element I_2 . By proposition 4.2 in section 4.3.1, this system can not have a relative degree through any dynamic extension. In fact, from the outputs point of view, the two control inputs are misplaced so that their effects are not separable. Thus, the independent tracking controls of both y_1 and y_2 can not be performed at the same time.

From the above discussion, it appears that unless the system dynamics is modified or augmented to reroute the shortest causal paths, not only the decoupling matrix will always be singular but also the controller design will be very difficult. One simple solution for this problem is redesigning the input locations. For example, if u_2 is

applied on element I_3 , two non-overlapping shortest causal paths can be selected as shown in Figure 4.13. Therefore, the relative degree and zero dynamics of this system can be defined. The independent controls of both outputs become possible. Note that this solution is not easy to see if only system equations are used.

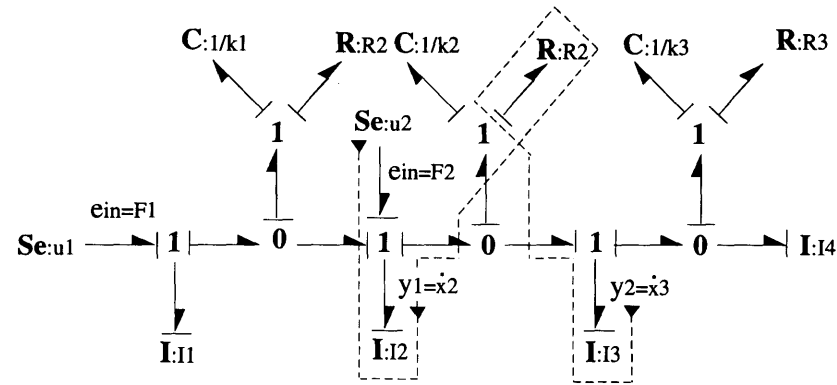


Figure 4.11: The shortest causal paths of a MIMO model.

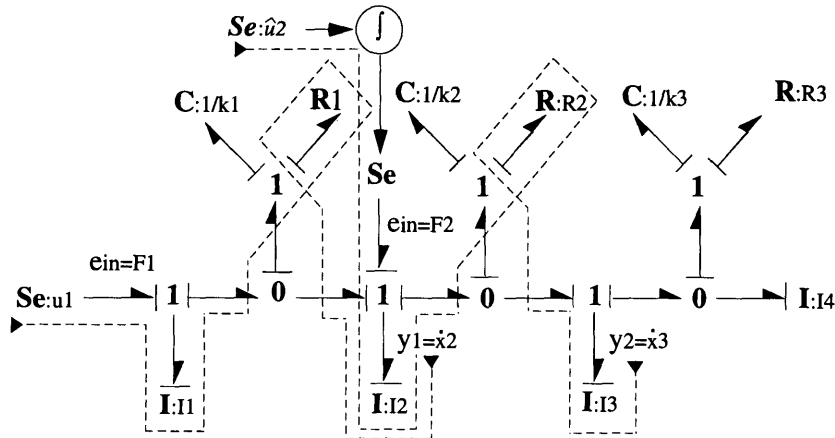


Figure 4.12: The shortest causal paths of a MIMO model by dynamic extension.

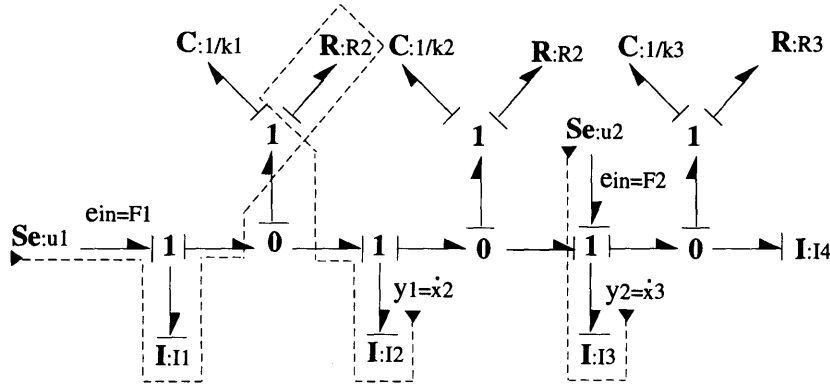


Figure 4.13: An alternative input-output configuration design.

4.6 Conclusion

In this chapter, the proposed Zero Dynamics Identification Procedure (ZDIP) presented in chapter 3 is extended for MIMO systems. It is shown that for MIMO systems, an essential issue of system design is the existence of the vector relative degree. If a system has a vector relative degree, its zero dynamics can be identified by a straightforward extension of the SISO ZDIP procedure. If a system does not have a vector relative degree, a dynamic extension procedure may be used to fix the structure. Then, the zero dynamics can still be identified in a similar manner. It is also shown that if the input-output configurations are ill-designed, not only the relative degrees do not exist, but also the zero dynamics can not be reasonably defined. In that case, separate tracking controls for the outputs are impossible. As a result, a useful guideline is provided for the design of the input-output configurations as well as the zero dynamics of MIMO systems.

Decomposition of Linear Dynamics in the Physical Domain and Eigenvalue Estimations

5.1 Introduction

The eigenvalues of a linear system is the most essential dynamic feature in the design considerations. From the system design point of view, it dominates the open-loop system behaviors and contributes to the performance limits of feedback systems. In an analysis process, the numerical eigenvalues can be easily computed by existing software programs. However, for the purpose of design, these numerical values do not indicate any possible improvement toward better system performance unless by a huge number of trial-and-error iterations. Therefore, it is important to build the direct relations between the component characteristics and the system eigenvalues in order to perform a systematic design.

It is known that the symbolic solutions for the eigenvalues of high order systems are not available. Even if the exact solutions exist, they may be too complicated, and therefore do not point out useful design directions. So instead of using the exact solutions, the approximations or the bounds of the eigenvalues may be more feasible. If they can be found by simple computations, the influences of the system components would be shown effectively. In the literature, many efforts has been made to find the numerical bounds of the eigenvalues to save the computation time for a large system [4]. Also, a variety of matrix theories have been proposed to find the bounds of

the eigenvalues in terms of the matrix components [16]. [51] proposes a method to obtain the formulas of the eigenvalues for a class of systems with uniform parameters¹. However, in many design cases, these approaches still do not provide a satisfactory result.

In this chapter, the difficulty of directly using the existing approaches is examined. Several decomposition procedures are proposed to improve the results. These procedures identify the physical components which contribute most to certain group of eigenvalues. By using the available matrix theories or other existing approaches, the bounds of each eigenvalue group can be represented in terms of the component characteristics. These bounds will then facilitate the design of physical systems so that the eigenvalues are approximately at the desired locations.

In section 2, currently available methods are examined. Section 3 describes the decomposition procedure for fast-slow dynamics. Section 4 shows the decomposition procedure for high-low frequency oscillation modes. Section 5 shows the decomposition procedure for the heavily-damped modes and the lightly-damped modes. The eigenvalue estimations for general systems are discussed in section 6. Several design examples are shown in section 7. The conclusion is given in section 8.

5.2 Currently Available Methods

The currently available methods for eigenvalues estimation can be divided into two categories. One is using the matrix theories such as the Gersgorin's theorem and its many versions of extensions [16]. These theories give a simple estimation of the bounds in terms of the matrix components. For example, given a complex $n \times n$

¹A system has uniform parameters if all the inertance elements, capacitance elements and dissipation elements in the system have the same parameters respectively.

matrix \mathbf{A} ,

$$\begin{bmatrix} a_{11} & a_{12} & \dots & \dots & a_{1n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{j1} & a_{j2} & \dots & \dots & a_{jn} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & \dots & a_{nn} \end{bmatrix} \quad (5.1)$$

a set of Gersgorin discs $D_j(\mathbf{A})$ can be formed on the S-plane by choosing the diagonal terms as the center and the absolute sum of the off-diagonal terms in each row as the radius,

$$D_j(\mathbf{A}) = \left\{ z \in \mathbf{C} : |z - a_{jj}| \leq \sum_{j \neq \ell} |a_{j\ell}| \right\}, \quad \text{for } j = 1, 2, \dots, n \quad (5.2)$$

The theorem proves that each eigenvalue of \mathbf{A} lies in some Gersgorin's disc of \mathbf{A} .

Although these theories are easy to apply, they can not be directly used for system design. One obvious reason is that since the physical system parameters are real numbers, the \mathbf{A} matrices of the state equations are real. In this case, the center of the Gersgorin's discs will be real. If the system possesses complex eigenvalues, the radii of some discs have to be very large to include the eigenvalues in the discs. Therefore, the bounds would be too far from the eigenvalues. To solve this problem, the \mathbf{A} matrices need be pre-conditioned so that the diagonal terms contain more system characteristics [11]. However, the purpose of this pre-conditioning procedure is different from those numerical procedures which are used to save the computation time in the finite element applications. Since the purpose is for system design, a symbolic matrix description is required after the pre-conditioning. Therefore, this procedure needs to be carefully designed so that a symbolic transformation can be carried out, yet the diagonal terms possess the most possible dynamic features of the system. The details of such a procedure will be discussed in section 5.6.2.

Another problem is that in some systems, the values of certain system parameters are much larger than the others. If the Gersgorin's theorem is applied to these

systems, some of the Gersgorin's discs will inevitably become large. For example, Figure 5.1 shows a simple R-C circuit. The state equations of this system are

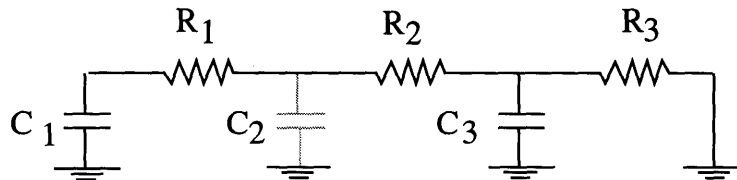


Figure 5.1: An R-C circuit.

$$\begin{pmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \dot{q}_3 \end{pmatrix} = \begin{bmatrix} \frac{-1}{C_1 R_1} & \frac{1}{C_2 R_1} & 0 \\ \frac{1}{C_1 R_1} & \frac{-(R_1 + R_2)}{C_2 (R_1 R_2)} & \frac{1}{C_3 R_2} \\ 0 & \frac{1}{C_2 R_2} & \frac{-(R_2 + R_3)}{C_3 (R_2 R_3)} \end{bmatrix} \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} = \mathbf{A} \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} \quad (5.3)$$

A set of symbolic bounds for the eigenvalues can be derived by Gersgorin's theorem. Suppose the parameters of this system are chosen as $C_1 = C_3 = 1$, $C_2 = 0.1$, $R_1 = R_2 = R_3 = 1$. The numerical \mathbf{A} matrix becomes

$$\mathbf{A} = \begin{bmatrix} -1 & 20 & 0 \\ 1 & -20 & 1 \\ 0 & 10 & -2 \end{bmatrix} \quad (5.4)$$

The numerical bounds are shown in Figure 5.2. From the passivity of this system and the fact that this system has no inertance elements, it can be concluded that the eigenvalues must be on the negative real axis. The bounds are still too large to provide any indication of the eigenvalue locations. For this type of systems, the eigenvalues usually can be separated into different groups which are far from each others. Therefore, unless the physical elements which are mostly responsible for each group of eigenvalues are identified and partitioned, no matter whether pre-conditioning procedures are applied, the Gersgorin's theorem can not obtain useful results.

The second category of approaches assume that the system has uniform parameters, In that case, it would be easier to obtain the eigenmodes of the physical systems.

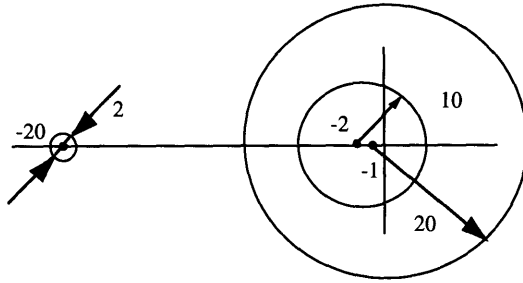


Figure 5.2: An estimation of eigenvalues from Gersgorin's theorem.

Then the eigenvalues can be easily calculated and the formulas can be formed. For example, Figure 5.3 shows the first oscillation mode a mass-spring system. Since the system has uniform parameters, there must be a node at exactly the center of the system. Therefore, the eigenvalue of this mode can be determined by either one of the subsystems separated by the node. Similarly, the highest frequency oscillation mode of the same system is shown in Figure 5.4. In this case, there is a node in the middle of each adjacent mass pair. The eigenvalue of each subsystem should be the same and equal to the eigenvalue of this mode. So the exact positions of the nodes can be easily identified and a formula for the eigenvalue of this mode can be obtained.



Figure 5.3: The first oscillation mode of a mass-spring system.

However, for general systems with non-uniform parameters, the only result this method can obtain is certain bounds of the eigenvalues. These bounds are obtained by forming systems with the uniform parameters which generate the largest and smallest



Figure 5.4: The highest frequency oscillation mode of a mass-spring system.

possible eigenvalues. It is obvious that if a system has elements with very different values, these bounds would be very big.

Therefore, to obtain meaningful bounds for the eigenvalues, the physical systems should be decomposed in a way that each subsystem represents a compact group of eigenvalues if it is possible. Then many techniques can be used to estimate the eigenvalues and get tighter bounds. Since these bounds are closely related to the characteristics of physical elements, they can be directly used in a design process.

In the following sections, three decomposition procedures are presented individually for certain categories of dynamic systems. Then the considerations for general systems follows.

5.3 Decomposition of Fast-Slow Dynamics

One common technique in the application of eigenvalues is the decomposition of fast-slow dynamics. When a system contains fast and slow dynamics, it is well known that the slow dynamics dominate the system behavior. Therefore, the eigenvalues corresponding to the fast dynamics usually can be safely ignored in the analysis. However, this process does not involve the identification of physical elements which contribute to the fast and slow dynamics. If a design task needs to modify the eigenvalues of the dominant dynamics, such a numerical decomposition does not help.

5.3.1 Singular perturbation theory

According to the singular perturbation theory [26], if a system has dynamics with different time scales, the state equations can be decomposed as follows.

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} \Rightarrow \begin{pmatrix} \dot{\mathbf{x}}_1 \\ \epsilon\dot{\mathbf{x}}_2 \end{pmatrix} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix}, \quad \epsilon \ll 1 \quad (5.5)$$

, where \mathbf{x}_1 is the state vector of the slow dynamics, \mathbf{x}_2 is the state vector of the fast dynamics, ϵ is a normalization factor. In this representation, all the components in the \mathbf{A} matrix have the values with the same order of magnitude.

From the fast dynamics point of view, the states of the slow dynamics are quasi-static. Therefore, the fast dynamics can be represented as

$$\epsilon\dot{\mathbf{x}}_2 = \mathbf{A}_{22}\mathbf{x}_2 \quad (5.6)$$

Since the fast dynamics has much faster transient response, the slow dynamics will evolve with the states \mathbf{x}_2 at equilibrium status, i.e. $\dot{\mathbf{x}}_2 = 0$. So the slow dynamics can be derived as

$$\dot{\mathbf{x}}_1 = [\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}]\mathbf{x}_1 \quad (5.7)$$

With this approach, if the computation can be carried out with symbolic descriptions, the system elements which contribute to the fast and the slow dynamics can be identified individually. However, this method would fail to explore the influence of the system structures to the eigenvalue locations.

5.3.2 Decomposition in the physical domain

To include the system structure information, the decomposition should be carried out in the physical domain. Namely, this decomposition should be performed directly on a system model such as bond graph models. For a class of systems, this is particularly

easy. If the system contains only R , C elements or R , I elements, the eigenvalues will always be real. The elements which are involved with the fast dynamics or the slow dynamics can be identified by the local loop gains [45]. For example, a bond graph model of the system in Figure 5.1 is shown in Figure 5.5. In this model, the element

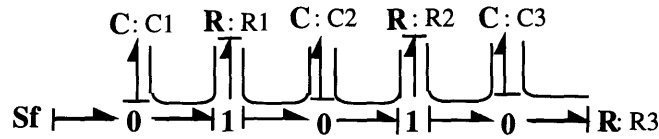


Figure 5.5: The bond graph model of an R-C system.

C_1 impose effort to the zero junction, then through the one junction to the element R_1 . The element R_1 impose flow to the one junction and through the zero junction to the element C_1 . Therefore, a causal loop is formed between these two elements. The loop gain of this path is $\frac{1}{R_1 C_1}$. Similarly, a causal loop is formed between the elements R_1 and C_2 , C_2 and R_2 , R_2 and C_3 , C_3 and R_3 . Suppose the element C_2 has a particularly small value, the loop gains $\frac{1}{R_1 C_2}$ and $\frac{1}{R_2 C_2}$ will become much larger than others. This means that the energy stored in the capacitance C_2 will be dissipated by R_1 and R_2 very quickly. Therefore, the elements R_1 , C_2 and R_2 along with the junctions which the causal loops pass through represent the fast dynamics as shown in Figure 5.6. Once the fast dynamics reaches its equilibrium status $\dot{q}_2 = 0$, the element C_2 plays no role in the slow dynamics. The condition $\dot{q}_2 = 0$ can be imposed by replacing the element C_2 with a flow source of 0 value. The slow dynamics can then be represented by the model shown in Figure 5.7.

Remark 1 : If the state equations are derived according to the models in Figure 5.6 and 5.7, they will be exactly the same as those derived from the perturbation theory. The inverse of the matrix \mathbf{A}_{22} in Eqn. (5.7) is automatically solved by the manipulation of the causality. With this approach, the physical elements and the

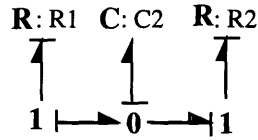


Figure 5.6: The bond graph model of the fast dynamics.

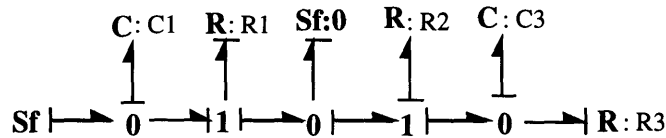


Figure 5.7: The bond graph model of the slow dynamics.

system structures which are responsible for the fast or the slow dynamics can be clearly identified.

Remark 2 : If this method is applied to electrical circuits, it is identical to the well-known "open-circuit", "closed-circuit" manipulations. For example, when a local loop is isolated, the subsystems which do not belong to this loop are "open-circuited". As indicated in Figure 5.8, the local loop is formed when R_2 , C_2 are isolated. Also, when an inertance element is replaced by an effort source with zero value, it is equivalent to the case where the two ends of this element are "short-circuited" as shown in Figure 5.9.

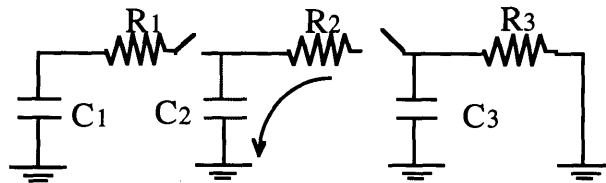


Figure 5.8: An isolated R-C loop.

Remark 3 : Using this method, even if the system becomes large, the number of loop

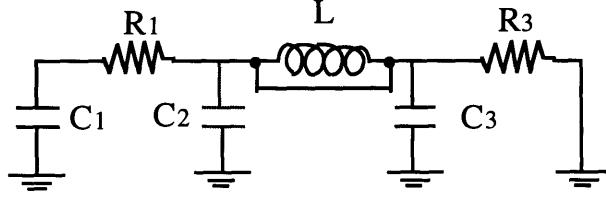


Figure 5.9: A short-circuited I element.

gains which need to be examined will not grow fast and become difficult to handle. For each energy storage element, only the elements which are directly causally related to it form local loops. Therefore, if the system contains N elements, the number of loop gains which need to be examined will be² kN and not N^2 or³ C_2^N .

5.3.3 A numerical example

To verify the results of this decomposition, the state equations corresponding to the fast and the slow dynamics are derived as follows according to Figure 5.6 and 5.7.

$$\dot{q}_2 = \frac{-(R_1 + R_2)}{C_2(R_1 R_2)} q_2 \quad (5.8)$$

$$\begin{pmatrix} \dot{q}_1 \\ \dot{q}_3 \end{pmatrix} = \begin{bmatrix} \frac{-1}{C_1(R_1+R_2)} & \frac{1}{C_3(R_1+R_2)} \\ \frac{1}{C_1(R_1+R_2)} & \frac{-(R_1+R_2+R_3)}{C_3 R_3(R_2+R_3)} \end{bmatrix} \begin{pmatrix} q_1 \\ q_3 \end{pmatrix} \quad (5.9)$$

If the system parameters of this example in the previous section are applied, the numerical \mathbf{A} matrix of the slow dynamics becomes

$$\begin{bmatrix} -0.5 & 0.5 \\ 0.5 & -1.5 \end{bmatrix} \quad (5.10)$$

The eigenvalue of the fast dynamics is -20 . The bounds obtained by the Gersgorin's theorem is shown in Figure 5.10. This result shows that although the decomposed systems represent only the approximations of the eigenvalues, the bounds provide a

² k is a constant.

³This is the number of all the possible combinations of any two elements in the system.

good estimation on the influences of the physical elements. The symbolic bounds obtained from Eqn. (5.9) and the approximated eigenvalue obtained from Eqn. (5.8) can be directly used for system design.

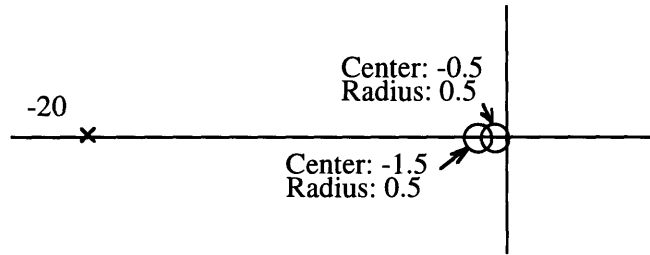


Figure 5.10: The bounds of the eigenvalues.

5.4 Decomposition of High-Low Frequency Oscillation Modes

If a system contains only energy storage elements, i.e. I elements or C elements, and no dissipative elements, the eigenvalues will be all on the imaginary axis, since such a system can only exhibit oscillations but not dissipations. In this case, the singular perturbation theory will fail to provide any conclusion. However, by the following auxiliary transformation, an $I-C$ network can be transformed into a fictitious $R-C$ or $R-I$ network. The singular perturbation theory and the decomposition procedure described in the previous section can then be applied to the fictitious systems if the system contains eigenvalues which are far apart. After transforming the system back to the original form, the subsystems which exhibit high-low frequency oscillation modes can be obtained respectively.

5.4.1 An auxiliary transformation

The state equations of an $I - C$ system can always be represented in the following general form if they are derived from a bond graph model.

$$\begin{pmatrix} \dot{\mathbf{e}} \\ \dot{\mathbf{f}} \end{pmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{C} \\ \mathbf{I} & \mathbf{0} \end{bmatrix} \begin{pmatrix} \mathbf{e} \\ \mathbf{f} \end{pmatrix} \quad (5.11)$$

, where \mathbf{e} is a state vector which represents the efforts associated with the C elements, \mathbf{f} is a state vector which represents the flows associated with the I elements, \mathbf{C} is a submatrix which contains the parameters associated with the C elements and \mathbf{I} is a submatrix which contains the parameters associated with the I elements. This set of state equations can also be represented as the following alternative forms.

$$\ddot{\mathbf{e}} = \mathbf{C}\mathbf{I}\mathbf{e} \quad \text{or} \quad \ddot{\mathbf{f}} = \mathbf{I}\mathbf{C}\mathbf{f} \quad (5.12)$$

Note that from this representation, it is clear that the nontrivial eigenvalues of Eqn. (5.11) will be the square roots of the eigenvalues of matrix $\mathbf{C}\mathbf{I}$ or matrix $\mathbf{I}\mathbf{C}$.

For the general $I - C$ systems discussed above, if all the I elements are replaced by R elements with the same parameters, the following equations can be derived from the bond graph model in a similar manner.

$$\begin{pmatrix} \dot{\mathbf{e}} \\ \dot{\mathbf{f}} \end{pmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{C} \\ \mathbf{R} & \mathbf{0} \end{bmatrix} \begin{pmatrix} \mathbf{e} \\ \mathbf{f} \end{pmatrix} \quad (5.13)$$

The state equations can then be represented as

$$\dot{\mathbf{e}} = \mathbf{C}\mathbf{R}\mathbf{e} \quad (5.14)$$

Note that the matrix $\mathbf{C}\mathbf{R}$ will be exactly the same as the matrix $\mathbf{C}\mathbf{I}$ of the original system.

Similarly, if all the C elements are replaced by R elements with the same parameters. the following equations can be derived from the bond graph model.

$$\begin{pmatrix} \dot{\mathbf{e}} \\ \dot{\mathbf{f}} \end{pmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{R} \\ \mathbf{I} & \mathbf{0} \end{bmatrix} \begin{pmatrix} \mathbf{e} \\ \mathbf{f} \end{pmatrix} \quad (5.15)$$

The state equations can be represented as

$$\dot{\mathbf{f}} = \mathbf{IRf} \quad (5.16)$$

Note that the matrix \mathbf{IR} will be exactly the same as the matrix \mathbf{IC} of the original system.

By the above derivations, it can be concluded that if the eigenvalues of matrix \mathbf{CR} (or \mathbf{CI}) can be separated into two groups which represent fast-slow dynamics, the eigenvalues of Eqn. (5.11) can be separated into two groups which are responsible for the high-low oscillation modes. Similarly, if the eigenvalues of matrix \mathbf{IR} (or \mathbf{IC}) can be separated into two groups which represent fast-slow dynamics, the eigenvalues of Eqn. (5.11) can be separated into two groups which are responsible for the high-low oscillation modes. Therefore, an auxiliary transformation can be defined so that the singular perturbation theory can be used to justify the decomposition of oscillation modes. This transformation procedure simply replaces all the I or C elements in an $I - C$ system by R elements with the same parameters. The effect of such a transformation can be visualized by Figure 5.11. This transformation brings the eigenvalues of the original systems from the imaginary axis to the real axis by a one-on-one mapping. Note that in the actual implementations, such a derivation or transformation is not necessary. Since this transformation is always possible, the application of the decomposition procedure discussed in the previous section is extended to $I - C$ systems without any modification.

5.4.2 Physical interpretations

By applying the decomposition procedure, the elements and structures which are responsible for the high-low frequency oscillation modes can be identified in a systematic way. These results have their physical interpretations, which are sometimes observed

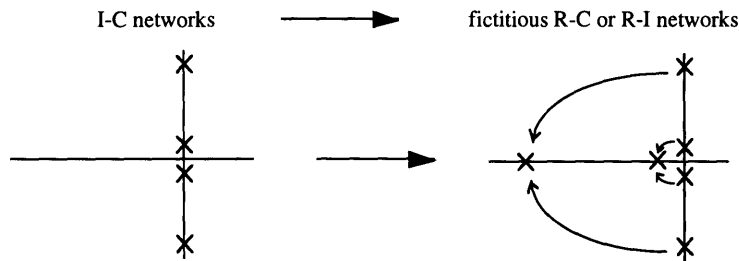


Figure 5.11: The effects of the auxiliary transformation.

by experienced system designers. An important advantage of the proposed systematic procedure is that it provides an easy solution for programming, while an inspection procedure based on the designer's intuition is difficult to do so. In the following, two examples are used to illustrate the physical interpretations of the decomposition results.

Figure 5.12 shows a simple cascaded mass-spring system. In the first case, suppose that the element C_2 has a much smaller value, i.e. this spring is much stiffer than others, and the other elements have values with the same order of magnitude. By examining the local loop gains, it will be found that the loop gain associated with elements C_2, I_1 ($\frac{k_2}{m_1}$) and the one associated with C_2, I_2 ($\frac{k_2}{m_2}$) are much larger than others ($\frac{k_1}{m_1}, \frac{k_3}{m_2}, \frac{k_3}{m_3}$). The decomposition procedure concludes that the subsystem shown in Figure 5.13 represents the high frequency oscillation mode. Also, the subsystem shown in Figure 5.14 represents the low frequency oscillation modes. In this subsystem, since the elements I_1 and I_2 are directly causally connected, i.e. there is a causal loop between these two elements, according to the results in Chapter 2, they can be grouped and represented by an equivalent I element as shown in Figure 5.15. The physical interpretation of this decomposition is that in the high frequency

subsystem representing the low frequency oscillation mode is derived by replacing I_2 with a flow source of zero value as shown in Figure 5.17. In this subsystem, since the elements C_2 and C_3 are directly causally connected, they can be grouped into an equivalent C element as shown in Figure 5.18. The physical interpretation of this decomposition is that in the high frequency oscillation mode, the large inertance elements behave like rigid boundaries. On the other hand, in the low frequency oscillation mode, the small mass almost has no effect on the dynamics. Therefore, it does not appear in the model.

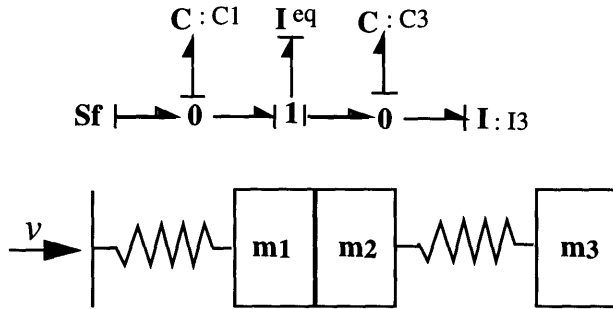


Figure 5.15: Case 1: low frequency oscillation modes.



Figure 5.16: Case 2: high frequency oscillation modes.

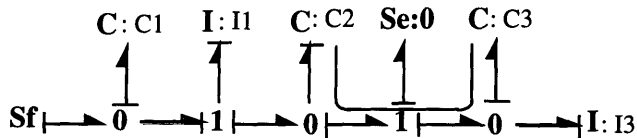


Figure 5.17: Case 2: low frequency oscillation modes.

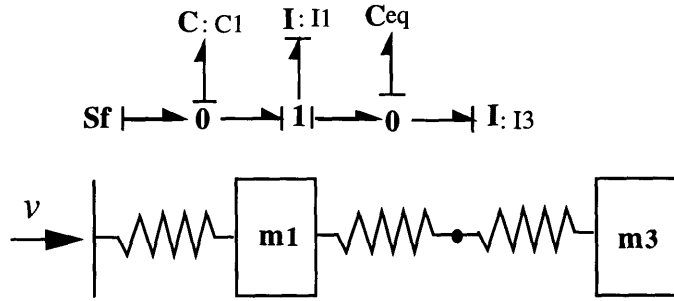


Figure 5.18: Case 2: low frequency oscillation modes.

5.5 Decomposition of Heavily-damped and Lightly-damped Dynamics

In the previous sections, it is shown that for $R - C$, $R - I$ or $I - C$ networks, a simple procedure can be employed to decompose the physical systems according to their eigenvalue distributions. For general systems, I , R , C elements will present at the same time. However, under certain assumptions, the proposed decomposition procedure can be reasonably applied. For example, if a system contains very little dissipation, the eigenvalues will be very close to the imaginary axis. Therefore, the system can be legitimately considered as an $I - C$ network. On the other hand, if a system contains very large dissipation everywhere, the eigenvalues will be separated into two groups on the real axis. One group represents faster dynamics. The physical system behaves like an $I - R$ network under the corresponding eigenmodes. The C elements contribute very little to these modes. Another group will be close to the origin. The physical system behaves like an $R - C$ network under the corresponding eigenmodes. Since this dynamics is slow, the I elements have no obvious effect⁴. For each group, the proposed procedure can be applied if a decomposition is possible.

⁴For example, in the case of a simple second order system $m\ddot{x} + R\dot{x} + kx = 0$, the two roots approach to $\frac{R}{m}$ and $\frac{k}{R}$ when R gets large.

In this section, the decomposition procedure will be extended for the systems which contain both light and heavy dissipations. The eigenvalue distribution of such systems is shown in Figure 5.19. The eigenvalues will be either close to the imaginary axis or to the real axis. The purpose of this decomposition is to identify the subsystems which are responsible for these two groups of eigenvalues

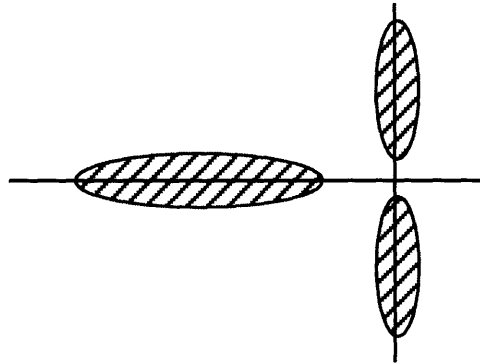


Figure 5.19: The eigenvalue distribution of the systems with both light and heavy dissipations.

5.5.1 The decomposition procedure

The decomposition procedure is still based on the local loop gains as discussed before. The $I - R$ or $C - R$ loop gains represent the energy dissipation rates in the local loops for the corresponding energy storage elements. On the other hand, the square roots of the $I - C$ loop gains represent the energy exchange rates in the local loops. Therefore, for each energy storage element, instead of a single loop gain, the local damping ratio becomes a dominant factor. For each directly causally related $I - C$ pair in the system (with R elements causally connected to either the I or C or both elements), the local damping ratios are determined as $\frac{G_{RC}}{2\sqrt{G_{IC}}}$ and $\frac{G_{RI}}{2\sqrt{G_{IC}}}$ for the C and the I elements respectively, where G_{IC} is the $I - C$ loop gain, G_{RC} is the sum of the $R - C$ loop gains and G_{RI} is the sum of the $R - I$ loop gains. This is the same way the damping ratio of a standard second order system is determined. In the following, a

procedure is presented to decompose the physical system into a subsystem H , which represents the heavily-damped eigenmodes and a subsystem L , which represents the lightly-damped eigenmodes.

Decomposition Procedure 1:

(1.1) *Replace all the C elements by flow sources with zero value, identify the remaining $R - I$ pairs which are directly causally related. Denote these $R - I$ elements and the involved junctions as part of the subsystem H .*

(1.2) *Restore the C elements which are replaced in step (1.1), identify the C elements which are directly causally related to the above I elements. If $\sqrt{G_{IC}} \gg G_{IR}$, replace the C elements by flow sources with zero value. Denote these flow sources as part of the subsystem H . If $\sqrt{G_{IC}} \ll G_{IR}$, ignore the identified C elements.*

(1.3) *Identify the I elements which become dependent due to the causalities imposed by the above sources. Denote these I elements as part of the subsystem H .*

(1.4) *Replace all the I elements by effort sources with zero value, identify the remaining $R - C$ pairs which are directly causally related. Denote these $R - C$ elements and the involved junctions as part of the subsystem H .*

(1.5) *Restore the I elements which are replaced in step (1.1), identify the I elements which are directly causally related to the above C elements. If $\sqrt{G_{IC}} \gg G_{RC}$, replace the I elements by effort sources with zero value. Denote these effort sources as part of the subsystem H . If $\sqrt{G_{IC}} \ll G_{RC}$, ignore the identified I elements.*

(1.6) *Identify the C elements which become dependent due to the causalities imposed by the above sources. Denote these C elements as part of the subsystem H .*

(1.7) *Identify the resistances which are involved in heavily-damped local loops (loops with large local damping ratio). Denote these R elements and the involved $I - C$ pairs, junctions as part of the subsystem H .*

(1.8) *Identify the C elements which are not involved in step (1.7), but are directly*

causally related to the above I elements. If $\sqrt{G_{IC}} \gg G_{IR}$, replace the C elements by flow sources with zero value. Denote these flow sources and the involved junctions as part of the subsystem H . If $\sqrt{G_{IC}} \ll G_{IR}$, ignore the identified C elements.

(1.9) Identify the I elements which are not involved in step (1.7), but are directly causally related to the above C elements. If $\sqrt{G_{IC}} \gg G_{RC}$, replace the I elements by effort sources with zero value. Denote these effort sources as part of the subsystem H . If $\sqrt{G_{IC}} \ll G_{RC}$, ignore the identified I elements.

(1.10) Remove the elements which are not denoted as part of the subsystem H . The remaining subsystem is the heavily-damped subsystem H .

Decomposition Procedure 2:

(2.1) Identify the $I - C$ pairs which are involved in lightly-damped local loops (loops with small local damping ratio), denote these $I - C$ elements as part of the subsystem L .

(2.2) Identify the R elements which are not involved in step 1, but are directly causally related to the above I or C elements. If $\sqrt{G_{IC}} \ll G_{RI}$ or $\sqrt{G_{IC}} \ll G_{RC}$, replace the resistive R elements by flow sources with zero value, and conductive R elements by effort sources with zero value, Denote these sources as part of the subsystem L .

(2.3) Identify the energy storage elements which become dependent due to the causalities imposed by the above sources. Denote these elements as part of the subsystem L .

(2.4) Remove the elements which are not denoted as part of the subsystem L . The remaining subsystem is the lightly-damped subsystem L .

Remark 1 : In procedure 1, step (1.1) identifies the R elements and the I elements which are surely responsible for the heavily-damped modes since even if all the capacitances are disabled, these elements still has dynamics. Step (1.2) and step (1.3) identify the I elements which are involved in the heavily-damped modes by the power

transmission through $I - C$ loops. Step (1.4) to step (1.6) repeat the same procedure for the $R - C$ elements. Step (1.7) includes the over-damped subsystems. Step (1.8) and (1.9) identify the I or C elements which are involved in the heavily-damped modes by the power transmission through other $I - C$ loops.

Remark 2 : In procedure 2, step (2.1) includes the lightly-damped subsystems. Step (2.2) and (2.3) identify the I or C elements which are involved in the lightly-damped modes by the power transmission through other $I - R$ loops.

Remark 3 : As will be shown by an example in section 5.7.2, if two elements are causally connected by multiple paths, one must notice that the effective loop gain is different from the sum of the loop gains corresponding to each individual path. In the latter case, the coupling between the multiple paths is not taken into account. Other than this, the use of the proposed method is not changed.

Example:

Consider a system shown in Figure 5.20. The bond graph model is shown in Figure 5.21. By applying the steps (1.1) and (1.4), it is found that no remaining $R - I$ or $R - C$ pairs are directly causally related. Therefore, only step (1.7) to (1.9) needs to be considered for procedure 1.

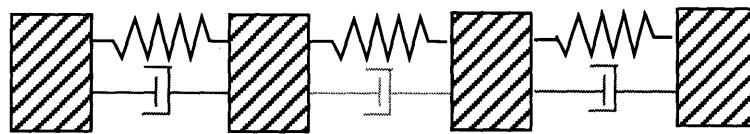


Figure 5.20: A simple mass-damper-spring system.

Suppose that in this system, R_2 has a large value, R_1 and R_2 are very small, the other elements has the values with the same order of magnitude. In this case, the $I - C$ pairs which are causally related to the element R_2 are involved in the heavily-damped modes. Since the natural motion of the lightly-damped subsystems

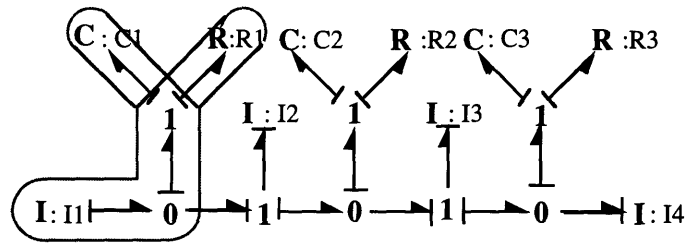


Figure 5.21: The corresponding bond graph model.

is oscillatory, their effects on the heavily-damped modes are negligible unless the loop gains of elements $I_2 - C_1$ and $I_3 - C_3$ are particularly large. The subsystem shown in Figure 5.22 represents the heavily-damped dynamics.

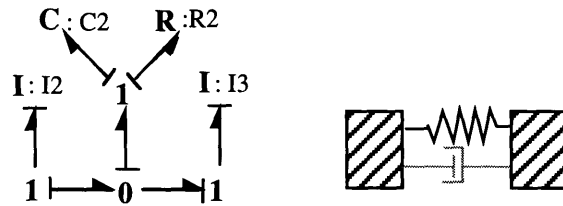


Figure 5.22: The bond graph model representing the heavily-damped modes.

On the other hand, when the system evolves under the oscillation modes, the subsystem associated with the resistance R_2 can not follow the motion easily since its natural motion is heavily-damped. As a result, this part of system behaves like a rigid mass in the oscillation modes. This constraint can be represented by replacing the resistance R_2 with a flow source with a zero value as shown in step (2.3). The corresponding model is shown in Figure 5.23. The model can also be represented as Figure 5.24, since the element C_2 plays no role in this system. According to step (2.2), if the subsystem associated with the resistance R_2 has a large damping ratio due to a small R_2 and an almost negligible C_2 , the model representing the oscillation modes becomes two separate subsystems as shown in Figure 5.25.

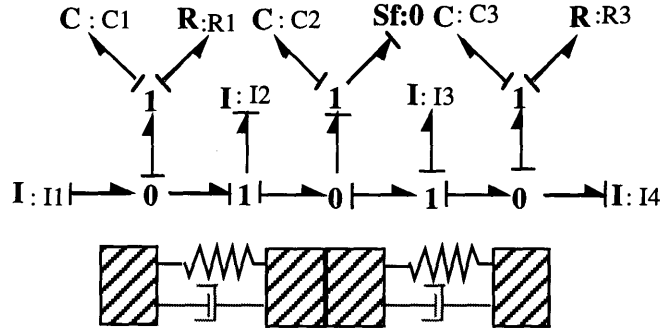


Figure 5.23: The bond graph model representing the lightly-damped modes.

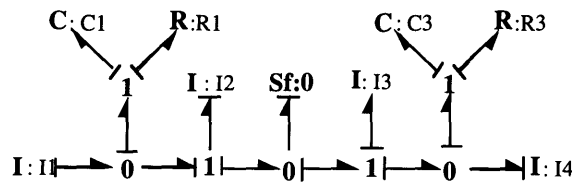


Figure 5.24: The equivalent bond graph model representing the lightly-damped modes.

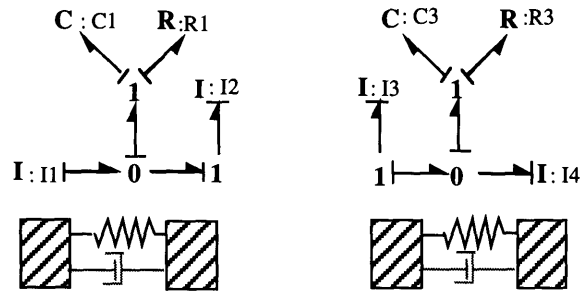


Figure 5.25: The bond graph model representing the lightly-damped modes.

5.5.2 A numerical example

As discussed before, the results from the decomposition procedure are approximations. To demonstrate the accuracy of such approximations, a numerical example is presented in the following. However, the purpose of the decomposition is not to obtain the numerical eigenvalues. The ultimate goal is to obtain the symbolic bounds of eigenvalues and to use them for system design.

Suppose that in the system of Figure 5.20, $m_1 = m_2 = m_3 = m_4 = 1$, and $k_1 = k_2 = k_3 = k_4 = 1$. Under this assumption, since the $I - C$ pairs all have

the same loop gains, there are only three different damping ratios $\zeta_1 = \frac{R_1}{\sqrt{m_1 k_1}}$ (from the local loops formed by R_1, I_1 and C_1, I_1 as indicated in Figure 5.21), or $\frac{R_1}{\sqrt{m_2 k_1}}$, $\zeta_2 = \frac{R_2}{\sqrt{m_2 k_2}} = \frac{R_2}{\sqrt{m_3 k_2}}$, and $\zeta_3 = \frac{R_3}{\sqrt{m_3 k_3}} = \frac{R_3}{\sqrt{m_4 k_3}}$. Figure 5.26 (a) shows the eigenvalue distribution of the case $\zeta_1 = \zeta_3 = 0.07$ and $\zeta_2 = 0.7$. Figure 5.26 (b) shows the eigenvalue distribution of the case $\zeta_1 = \zeta_3 = 0.07$ and $\zeta_2 = 1.4$.

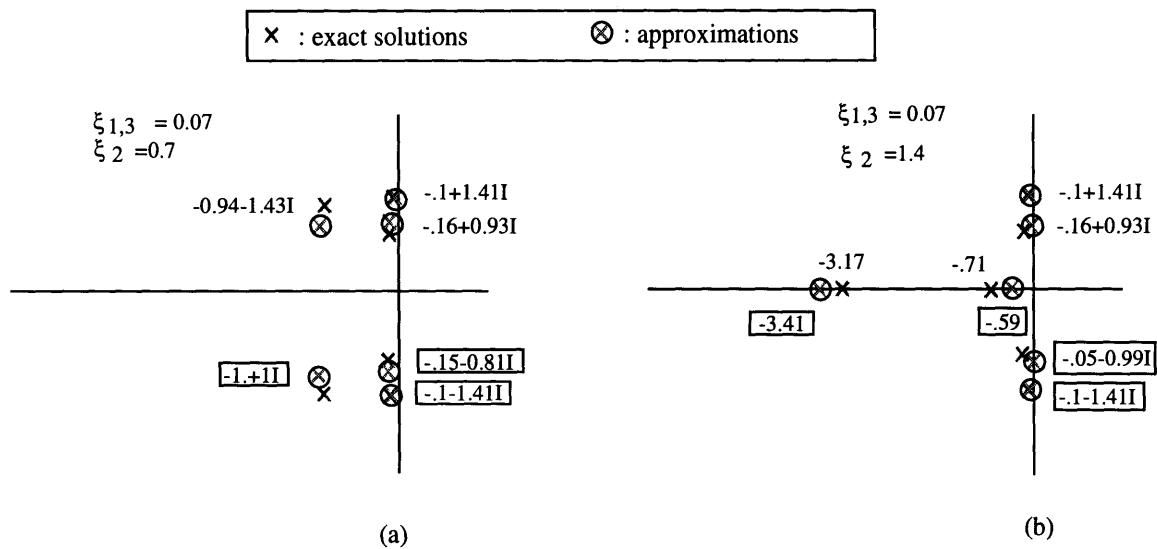


Figure 5.26: The decomposition results (a) $\zeta_2 = 0.7$. (b) $\zeta_2 = 1.4$.

Note that in the first case, the imaginary parts of the estimated eigenvalues with heavier damping ratio are off about 30%. This is because the damping ratio ζ_2 is not large enough. Therefore, the imaginary parts of these eigenvalues are close to the imaginary parts of the other eigenvalues. As a result, the coupling of these eigenvalues is not negligible. In case (b), the distribution of the eigenvalues shows the pattern as Figure 5.19. The estimations are much closer to the true eigenvalues.

5.6 Eigenvalue Estimation for General Systems

Based on the discussions in the previous sections, the proposed decomposition procedures are summarized in Figure 5.27. If all the local damping ratios are large, the system can be easily decomposed into a $R - C$ and a $R - I$ network⁵. The procedure described in section 5.3 can be applied to both networks if a decomposition is possible. On the other hand, if all the local damping ratios are small, the system can be treated as an $I - C$ network. The procedure in section 5.4 can be applied. If the system contains both large and small local damping ratios, the procedure in section 5.4 can be applied. Also, the decomposed heavily-damped subsystem can be decomposed again into $R - C$ and $R - I$ networks and be processed by the procedure in section 5.3.

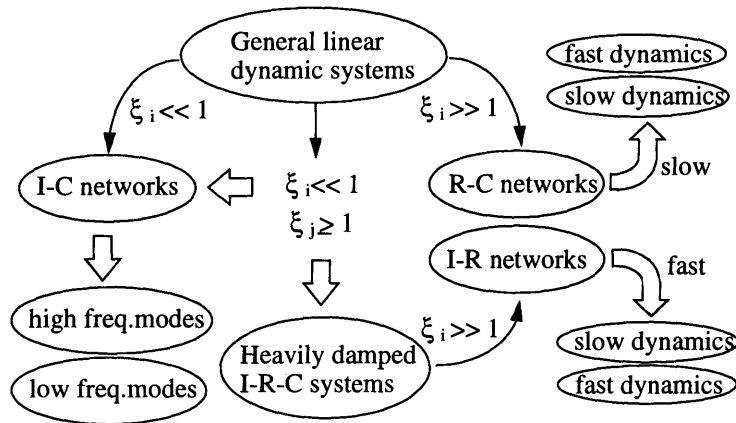


Figure 5.27: A summary of the proposed decomposition procedures.

By applying these procedures, it is possible that only a few eigenvalues out of a large system are extracted by the decomposition. In this case, whether the results provide useful information probably depends on the system itself. If the extracted

⁵This includes the case where only one type of energy storage elements appear in part of the system. In that case, the corresponding local damping ratios would be infinitely large.

eigenvalues are the dominant eigenvalues, the system's performance can be drastically improved by only modifying the characteristics of very few elements, which are strongly related to the dominant eigenvalues. On the contrary, if the extracted eigenvalues are far from the origin, the result would provide only a limited model reduction. As will be shown in the example of section 5.7.2, the result clearly indicates how to locate the oscillation poles and the real pole by choosing the parameters T_1 and T_2 . Even if this system is only a 3^{rd} order system, it is not easy to do the same by observing the \mathbf{A} matrix of the state equations.

5.6.1 Undecomposable systems

From the above summary, a missing link in dealing with the general systems is that if a system contains subsystems with local damping ratios in the middle of 0 and 1, none of the proposed procedures would apply. However, this does not mean that a decomposition is always impossible for the systems with moderate local damping ratios. For example, if the system eigenvalues have the distribution shown in Figure 5.28, it can be processed by a modified procedure from the results of the previous sections. On the other hand, it is almost impossible to provide a decomposition for the distribution in Figure 5.29. In this case, some physical elements may have

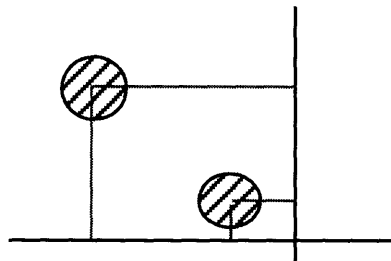


Figure 5.28: A decomposable distribution of eigenvalues.

significant influences in more than one group of eigenvalues. As a result, almost all the physical elements are important in every eigenmode. None of the subsystems

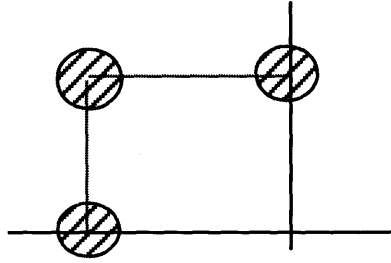


Figure 5.29: An undecomposable distribution of eigenvalues.

can be considered individually responsible for a certain eigenmode. This situation indicates that for such kind of systems, it is not easy to change the dynamic behavior by modifying only a few elements or subsystems. It may only be possible to move the whole group of eigenvalues by modifying the characteristics of all the energy storage elements or all the dissipative elements. As shown by the eigenvalue distribution of Figure 5.30, Butterworth type of filters are common practical examples of such undecomposable systems. For this category of systems, the best information that can be obtained for design may be the bounds of the eigenvalues in terms of the physical parameters. As described in section 5.2, the Gersgorin's theorem or similar matrix theories can not be directly applied to obtain meaningful bounds for design. In the following, a pre-conditioning procedure is proposed to provide a possible solution.

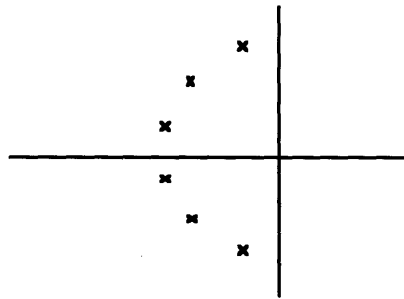


Figure 5.30: The eigenvalue distribution of a Butterworth type filter.

5.6.2 A pre-conditioning procedure

When deriving the state equations from bond graph models, the states are usually chosen as the power or the energy variables of the energy storage elements. In this case, each energy storage element is considered as a simplest subsystem. The system imposes a flow or effort to each of these subsystems and the subsystem reflects a effort or flow back to the system. The dynamics of such a subsystem is simply an integrator. By assembling these dynamics through the junction connections, a set of state equations can be obtained. For example, in Figure 5.31, each energy storage element provides an integrator. The state equations of this system will be in the following form.

$$\begin{pmatrix} \dot{\mathbf{e}} \\ \dot{\mathbf{f}} \end{pmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{C} \\ \mathbf{I} & \mathbf{R} \end{bmatrix} \begin{pmatrix} \mathbf{e} \\ \mathbf{f} \end{pmatrix} \quad (5.17)$$

As discussed in section 5.2, such an \mathbf{A} matrix do not satisfy our needs for finding

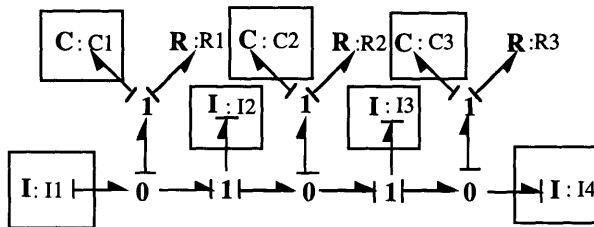


Figure 5.31: A bond graph model.

useful eigenvalue estimations. This is because the eigenvalue information scatters in the matrix, which does not fit the application of Gersgorin's theorem. Therefore, a pre-conditioning procedure is necessary to move the eigenvalue information to the diagonal terms as much as possible. However, for the purpose of design, the existing numerical procedures can not be used since a symbolic manipulation is required. Under this constraint, the tolerable complexity of the computation is very limited.

To solve this problem, instead of performing a transformation directly on the \mathbf{A} matrix, a different way of assembling the state equations provides a better solution.

For example, as an extension of the traditional method, the state equations can be derived by assembling the dynamics of the partitioned subsystems shown in Figure 5.32. For each subsystem, a set of state equations can be derived as

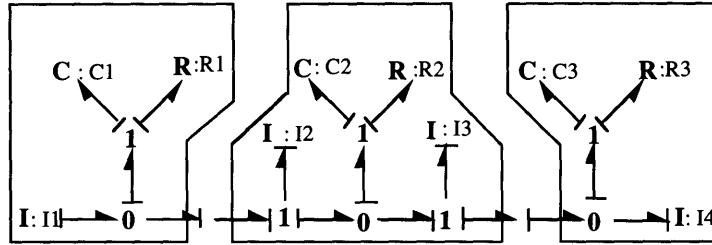


Figure 5.32: A simple partition.

$$\dot{\mathbf{x}}_i = \mathbf{A}_i \mathbf{x}_i + \mathbf{B}_i \mathbf{u}_i \quad (5.18)$$

$$\mathbf{y}_i = \mathbf{C}_i \mathbf{x}_i + \mathbf{D}_i \mathbf{u}_i \quad (5.19)$$

, where $i = 1, 2, 3$. Note that the inputs \mathbf{u}_i are from the outputs of other subsystems and the outputs \mathbf{y}_i are the inputs to other subsystems. Since the term $\mathbf{D}_i \mathbf{u}_i$ only reflects the input to the output, it can always be eliminated by combining its effects to the adjacent subsystems. If the state equations of these subsystems are assembled as the following, they will be exactly the same as Eqn (5.17).

$$\begin{pmatrix} \dot{\mathbf{x}}_1 \\ \dot{\mathbf{x}}_2 \\ \dot{\mathbf{x}}_3 \end{pmatrix} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{B}_1 \mathbf{C}_2 & \mathbf{0} \\ \mathbf{B}_2 \mathbf{C}_1 & \mathbf{A}_2 & \mathbf{B}_2 \mathbf{C}_3 \\ \mathbf{0} & \mathbf{B}_3 \mathbf{C}_2 & \mathbf{A}_3 \end{bmatrix} \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \end{pmatrix} \quad (5.20)$$

However, before the state equations are assembled, if a diagonalization is performed to each subsystem, the state equations will be in a more useful form. Suppose the state equations of each subsystem are diagonalized as follows

$$\dot{\mathbf{z}}_i = \mathbf{W}_i \mathbf{A}_i \mathbf{V}_i \mathbf{z}_i + \mathbf{W}_i \mathbf{B}_i \mathbf{u}_i \quad (5.21)$$

$$\mathbf{y}_i = \mathbf{C}_i \mathbf{V}_i \mathbf{z}_i \quad (5.22)$$

, where \mathbf{W}_i and \mathbf{V}_i are the left and the right eigenvector matrices of each subsystem.

The assembled state equations become

$$\begin{pmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \dot{z}_3 \end{pmatrix} = \begin{bmatrix} \mathbf{W}_1\mathbf{A}_1\mathbf{V}_1 & \mathbf{W}_1\mathbf{B}_1\mathbf{C}_2\mathbf{V}_2 & \mathbf{0} \\ \mathbf{W}_2\mathbf{B}_2\mathbf{C}_1\mathbf{V}_1 & \mathbf{W}_2\mathbf{A}_2\mathbf{V}_2 & \mathbf{W}_2\mathbf{B}_2\mathbf{C}_3\mathbf{V}_3 \\ \mathbf{0} & \mathbf{W}_3\mathbf{B}_3\mathbf{C}_2\mathbf{V}_2 & \mathbf{W}_3\mathbf{A}_3\mathbf{V}_3 \end{bmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} \quad (5.23)$$

In this set of state equations, the diagonal blocks contains the eigenvalues of each subsystems. The off-diagonal blocks represent the coupling effects between the subsystems. If the partitions are chosen so that the off-diagonal terms are small, while the symbolic computation for the diagonalization of the subsystems are possible, this method provides a useful transformation for general dynamic systems. Note that without the partitions in the physical domain, it is not easy to obtain Eqn. (5.23) directly from Eqn. (5.17) by symbolic manipulations. Therefore, this pre-conditioning procedure emphasizes the use of physical models. As a rule of thumb, the partitioned subsystems must be of low orders or with uniform parameters so that it is possible to obtain the eigenmodes in symbolic forms. In the cases where the characteristics of certain subsystems are not to be determined by the designer, i.e. these subsystems are considered with fixed parameters because of physical constraints or other reasons, numerical values will be used instead of symbolic forms. These subsystems with fixed numerical parameters should be grouped together whenever it is possible since the numerical diagonalization can always be performed.

5.7 Design Examples

To illustrate the proposed procedures, several design examples are presented in this section.

5.7.1 A mechanical structure

Figure 5.33 shows the schematic of a mechanical structure. The corresponding bond graph model is shown in Figure 5.34.

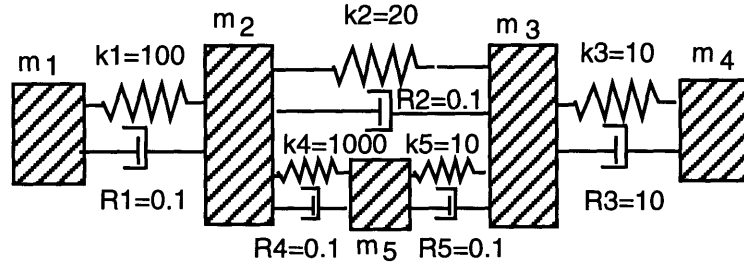


Figure 5.33: A mechanical structure.

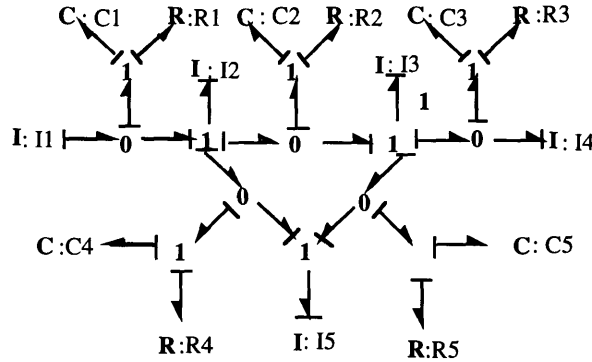


Figure 5.34: The corresponding bond graph model.

Assume that in this system, $m_1 = m_2 = m_3 = m_4 = m_5 = 1$ and the other parameters are as shown in Figure 5.33. First of all, the local damping ratio $\frac{R_3}{\sqrt{k_3 m_3}} = 3.16$ is much larger than any others (maximum 0.0316, minimum 0.0032). Therefore, according to the proposed procedure, the system can be decomposed into two subsystems representing the heavily-damped modes and the lightly-damped modes respectively as shown in Figure 5.35. Also, in the lightly-damped subsystem (as shown in Figure 5.35 (b)), the local loop gain $\frac{k_4}{m_5} = 1000$ is much larger than others (maximum

100, minimum 10). The lightly-damped subsystem can be further decomposed into high-low frequency oscillation modes as shown in Figure 5.36. Finally, as mentioned before, the two roots of a second order system $m\ddot{x} + R\dot{x} + kx = 0$ approach to $\frac{R}{m}$ and $\frac{k}{R}$ when R is large. Therefore, the heavily-damped subsystem can be further decomposed into fast-slow dynamics shown in Figure 5.37.

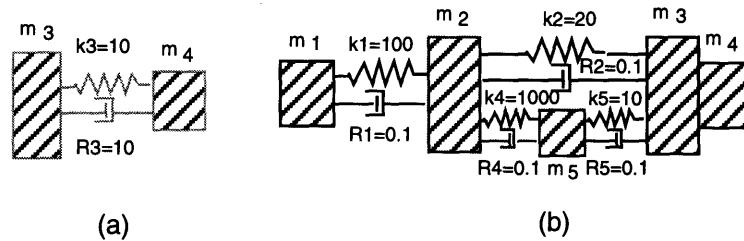


Figure 5.35: The decomposition of heavily-damped (a) and lightly-damped (b) modes.

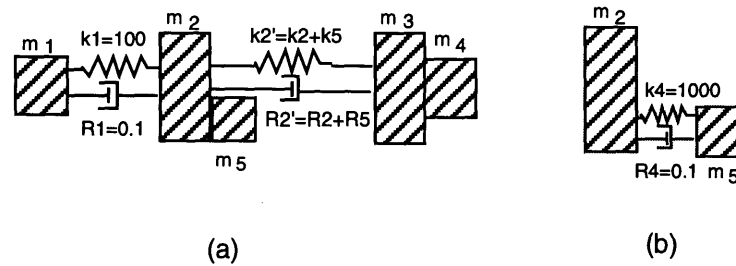


Figure 5.36: The decomposition of high (a) -low (b) frequency oscillation modes.

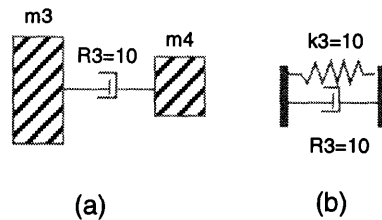


Figure 5.37: The decomposition of fast (a) -slow (b) dynamics.

The decomposed subsystems and the numerical eigenvalues are shown in Figure 5.38. Suppose that for a design task, the focus is only on the dominant dynamics

as circled in Figure 5.38. The considered physical system should include the low frequency oscillation modes: subsystem 2 and the slow dynamics: subsystem 4. As a result, the complex system in Figure 5.33 is simplified into the system in Figure 5.39. Since this simplified model reasonably approximate the behavior of the original system, the next step is to estimate the symbolic bounds of the eigenvalues.

As discussed in section 5.4, the state equations of an $I - C$ system can be written in the form of Eqn. (5.11). In this case, the A matrix of subsystem 2 in Figure 5.38 can be written as

$$\left[\begin{array}{cc|cc} 0 & 0 & \frac{1}{m_1} & \frac{-1}{m_2} & 0 \\ 0 & 0 & 0 & \frac{1}{m_2} & \frac{-1}{m_3} \\ \hline -k_1 & 0 & 0 & 0 & 0 \\ k_1 & -k_2' & 0 & 0 & 0 \\ 0 & k_2' & 0 & 0 & 0 \end{array} \right] \quad (5.24)$$

,where $m_2' = m_2 + m_5$, $m_3' = m_3 + m_4$, $k_2' = k_2 + k_5$ and $R_2' = R_2 + R_5$. The eigenvalues of this matrix will be the square roots of the eigenvalues of the matrix

$$\left[\begin{array}{ccc} \frac{1}{m_1} & \frac{-1}{m_2} & 0 \\ 0 & \frac{1}{m_2} & \frac{-1}{m_3} \end{array} \right] \left[\begin{array}{cc} -k_1 & 0 \\ k_1 & -k_2' \\ 0 & k_2' \end{array} \right] = \left[\begin{array}{cc} -\frac{k_1}{m_1} - \frac{k_1}{m_2} & \frac{k_2'}{m_2} \\ \frac{k_1}{m_2} & -\frac{k_2'}{m_2} - \frac{k_2'}{m_3} \end{array} \right] \quad (5.25)$$

By applying the Gersgorin's theorem, the bounds of the eigenvalues are shown in Figure 5.40. By these symbolic bounds, it can be concluded that the most efficient way to increase the lower bound of the oscillation frequency is to increase the value of element k_2 . On the other hand, the most efficient way to reduce the upperbound is by reducing the value of k_1 . However, when k_2' becomes too large and k_1 becomes too small, their roles exchange. Note that k_3 and R_3 have very little effects on the oscillation modes. They are mostly responsible for the over-damped mode.

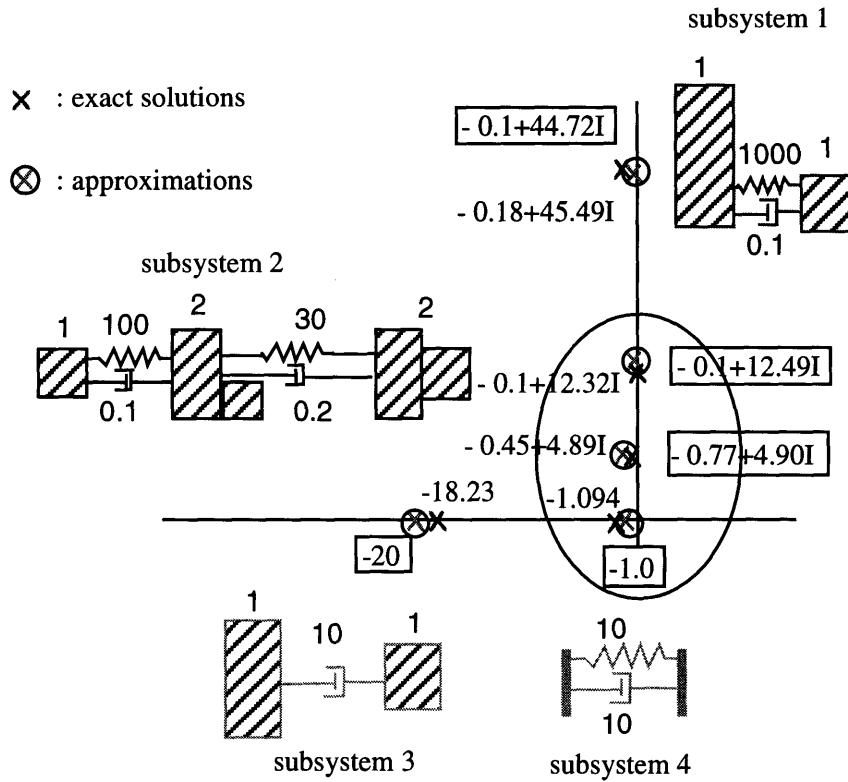


Figure 5.38: The estimated eigenvalues from the decomposed subsystems.

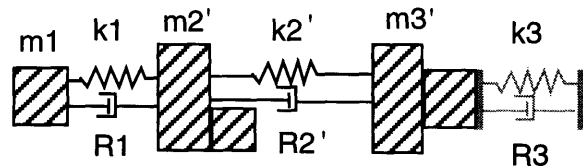


Figure 5.39: The physical systems representing the dominant dynamics.

5.7.2 An arm prosthesis design

In chapter 3, a simple arm prosthesis model is used to illustrate the design procedures concerning a system's zero dynamics. However, the eigenvalues of such a system has not been addressed. The bond graph model of this arm prosthesis is shown in Figure 5.41. The \mathbf{A} matrix of this system is shown below.

$$\mathbf{A} = \begin{bmatrix} 0 & -\frac{(\frac{1}{T_1} + \frac{1}{T_2})}{I_f} & \frac{1}{I_m T_1} \\ \frac{(\frac{1}{T_1} + \frac{1}{T_2})}{C_b} & -\frac{(R_m + \frac{R_i}{T_2} + R_f)}{I_f} & \frac{R_m}{I_m} \\ \frac{-1}{C_b T_1} & \frac{R_m}{I_f} & \frac{-R_m}{I_m} \end{bmatrix} \quad (5.26)$$

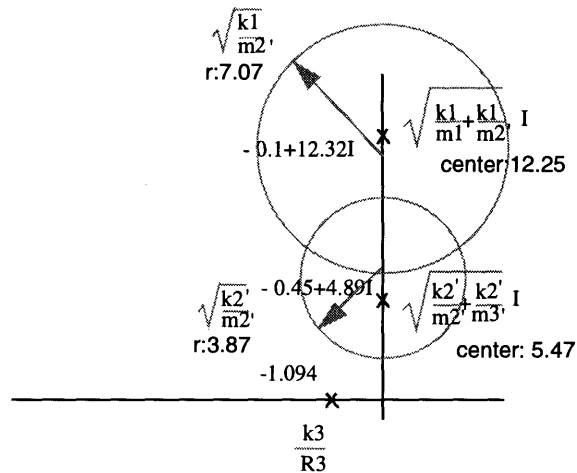


Figure 5.40: The oscillation modes for the dominant dynamics.

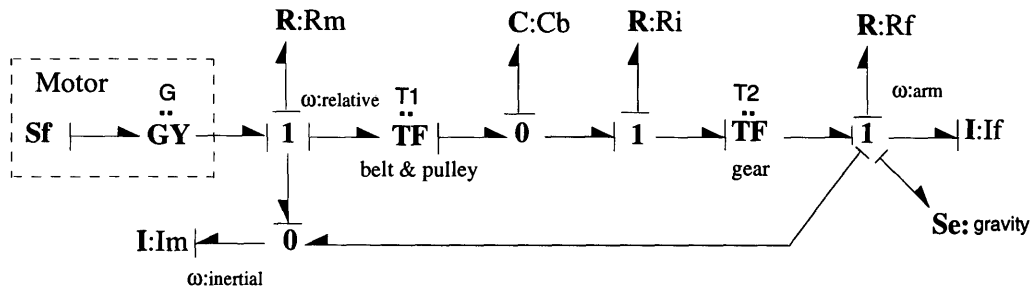


Figure 5.41: The bond graph model of an arm prosthesis.

The eigenvalue locations of this system are in a pattern shown in Figure 5.42, assuming that the dissipations are reasonably small. Although this is only a 3rd order system, there is no simple way to address the influence of the system parameters to the eigenvalues. On the other hand, the decomposition procedures discussed in this chapter can be employed to provide a useful insight about the relation between the system parameters/structures and the eigenvalues. First of all, by examining the causal relations, it can be found that the capacitance C_b has multiple causal paths leading to the inertia I_f . This does not change the fact that local loop gains serve as a guide line for decompositions. In this case, the effective loop gain associated with the multiple causal paths must be calculated. Such a loop gain associated with

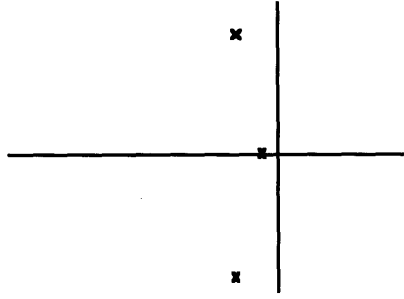


Figure 5.42: The eigenvalue distribution of the arm prosthesis.

C_b and I_f is $(\frac{1}{T_1} + \frac{1}{T_2})^2 \frac{1}{C_b I_f}$ as indicated by the dashed loop shown in Figure 5.43⁶.

Secondly, since the capacitance C_b is also causally related to the inertance I_m (the loop gain is $\frac{1}{T_1^2} \frac{1}{C_b I_m}$ as indicated by the dashed loop shown in Figure 5.44), the elements C_b , I_m , and I_f are all responsible for an oscillation eigenmode unless the dissipations are extremely large. If I_m is much larger than I_f or T_2 is particularly small, the C_b , I_f loop will dominate. In this case, the frequency of the oscillation mode would be close to $\sqrt{(\frac{1}{T_1} + \frac{1}{T_2})^2 \frac{1}{C_b I_f}}$. On the contrary, if I_f is particularly large, the frequency would be close to $\sqrt{\frac{1}{T_1^2} \frac{1}{C_b I_m}}$. On the other hand, if the loop gains concerning

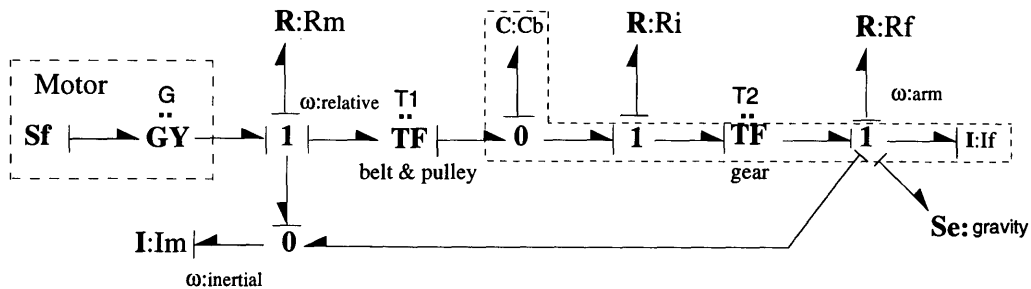


Figure 5.43: The local loop of C_b , I_f .

C_b , I_f and C_b , I_m are about the same order of magnitude, the coupling between these three elements can not be neglected. Figure 5.45 shows the bond graph model

⁶Note that this gain is not just the sum of the loop gains associated with each single causal path: $(\frac{1}{T_1^2} + \frac{1}{T_2^2}) \frac{1}{C_b I_f}$. The coupling between the two causal paths should be also considered.

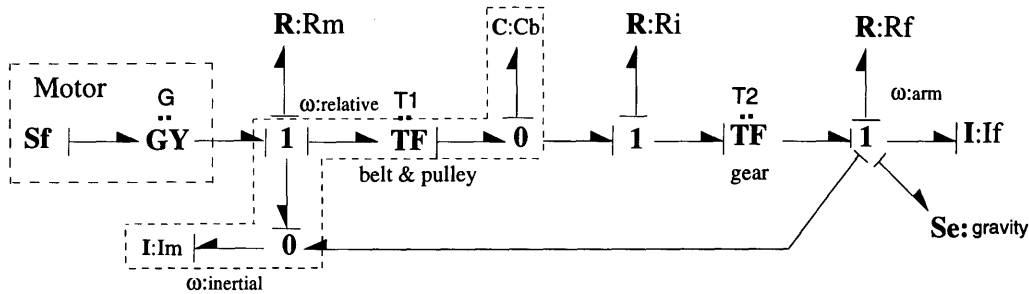


Figure 5.44: The local loop of C_b , I_m .

corresponding the pure oscillation eigenmode by removing the dissipations and the inputs. By the results in chapter 2, the subsystem enclosed by the dashed lines should be represented by an equivalent inertance. Figure 5.46 shows the revised model. From this model, the frequency of the oscillation mode should be $\sqrt{\frac{1}{T_1^2} \frac{1}{C_b I_m} + (\frac{1}{T_1} + \frac{1}{T_2})^2 \frac{1}{C_b I_f}}$. Since this is only an approximation, the exact value would be smaller due to the dissipations. However, this expression is already useful in designing the parameters to increase or decrease the frequency of the oscillation mode.

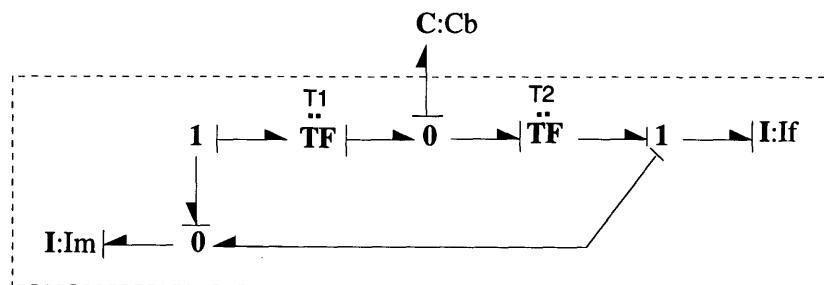


Figure 5.45: The bond graph model representing the pure oscillation eigenmode.

The system under consideration is a 3^{rd} order system. Therefore, the last eigenvalue must be a real number. In section 5.5, it is shown that a system with lightly-damped and heavily-damped modes can be effectively decomposed. A single real

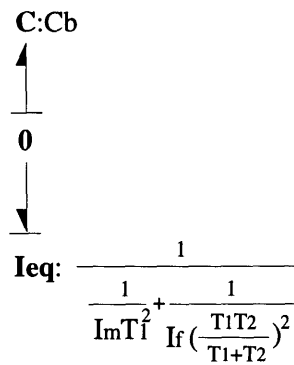


Figure 5.46: The equivalent bond graph model representing the pure oscillation eigenmode.

eigenvalue can be treated as an extreme case of heavily-damped modes. Also, in this arm prosthesis model, there is no dissipative elements directly causally connected to the capacitance C_b . So this real eigenvalue must be from the R and I elements. Assuming the all the R elements are reasonably small, by applying step (1.1) to (1.3) of the procedure in section 5.5, this real mode can be represented by replacing the capacitance using a flow source with zero value. The model is shown in Figure 5.47. From the causality, it can be concluded that this system has only one eigenvalue. The model can be simplified step by step as shown from Figure 5.48 to Figure 5.51. Finally, from Figure 5.51, the eigenvalue can be obtained as $\frac{R_m + R_i T_1^2 + R_f T_1^2 T_2^2}{I_m (1 + T_1 T_2)^2 + I_f T_1^2 T_2^2}$.

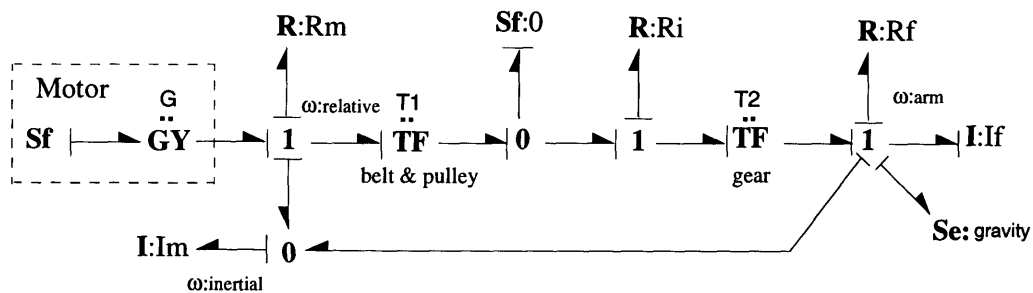


Figure 5.47: The bond graph model corresponding to the real eigenvalue.

To evaluate the accuracy of the above approximations, several numerical examples

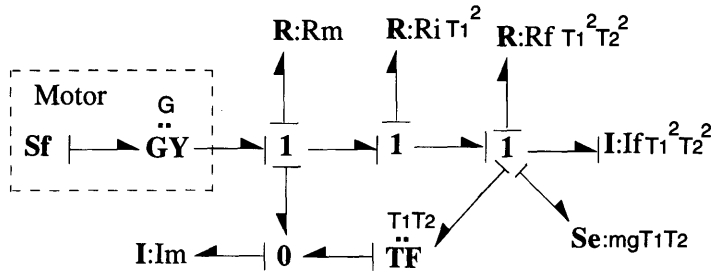


Figure 5.48: A simplified bond graph model.

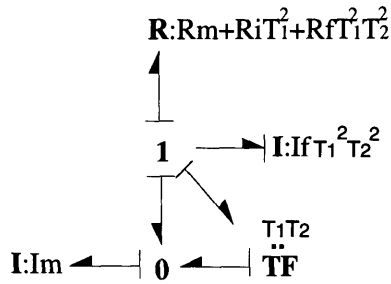


Figure 5.49: A simplified bond graph model.

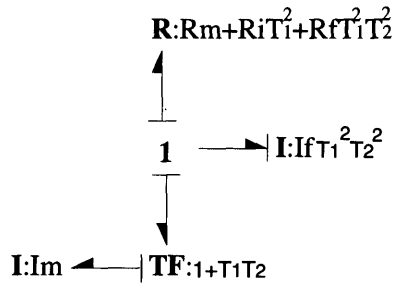


Figure 5.50: A simplified bond graph model.

are presented in Table 5.1. In these examples, the errors of the oscillation frequencies increase when the local damping ratios increase. However, if the values of the dissipative elements are held the same, the approximated oscillation frequencies are roughly proportional to the true frequencies. On the other hand, the approximated

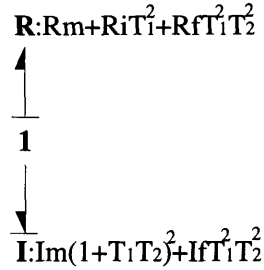


Figure 5.51: The bond graph model representing the real eigenmode.

real eigenvalues are very close to the true values when the values of R 's are within a range. In fact, if only the values of this system satisfy the decomposition conditions, the decomposed models give very approximations.

Note that the purpose of such decompositions in this example is not to obtain the eigenvalues with very precise numerical values or to save the computation efforts since there are other much more efficient ways to do so. The importance of this result is that the influences of the element characteristics such as the transformer parameters T_1 and T_2 to the eigenvalues are effectively shown by the obtained eigenvalue expressions. The result shows that T_1 and T_2 are coupled with other elements in particular ways for different groups of eigenvalues. Also, it is shown that C_b has very little influence to the eigenvalue on the real axis. These are very useful guidelines for the design of this arm prosthesis system, which can not be obtained by other approaches.

5.8 Conclusion

In this chapter, several decomposition procedures are proposed to identify the physical elements or structures which are responsible for separate groups of eigenvalues. For a category of systems, the bounds of the estimated eigenvalues are improved by the decomposition results. Since these bounds are represented by the physical parameters,

I_f	I_m	C_b	T_1	T_2	R_m	R_i	R_f	true eigenvalues	approximated eigenvalues
0.2	0.1	0.05	2	1.5	0.1	0.1	0.2	-1.022839605	$\pm 38.87301264j$
								$\pm 38.85747217j$	-0.6765430139
0.2	0.1	0.05	2	1.5	0.4	0.6	0.8	-4.192776386	$\pm 38.87301264j$
								$\pm 38.60241316j$	-2.947780567
0.2	0.1	1.0	2	1.5	0.4	0.6	0.8	-4.107743295	$\pm 8.692269875j$
								$\pm 7.375683411j$	-3.117846751
1.0	1.0	1.0	2	1.5	0.4	0.6	0.8	-0.7302916315	$\pm 3.333333333j$
								$\pm 3.226659773j$	-0.4060834040

Table 5.1: Numerical evaluations of the approximated eigenmode models

they can directly contribute to the design of physical systems. For general systems, the decomposition procedures do not always provide a satisfactory result. In this case, a pre-conditioning procedure along with the existing matrix theories may be applied to find the bounds of the eigenvalues. Several examples are presented to illustrate the use of the proposed procedures.

Chapter 6

Conclusion and Recommendations

In this thesis, three main topics are presented to demonstrate the use of the proposed structural analysis procedure. All three topics rely on the causalities and the junction structures in the bond graph representation to explore the structural information. The results provide useful insights for design and analysis of dynamic systems, which are impossible to obtain by using system state equations or transfer functions.

The first topic addresses the problem of excess states and their influences to the system analysis procedures. The excess states usually exist in certain over-constrained linear junction structures. In these models, although the representations are legitimate in terms of physical meaning, the resultant excess states cause pitfalls in the inspection of system properties. It is found that by using the explicit field representations, such ambiguities can be eliminated. Based on this approach, a set of model revision procedures are developed to eliminate the excess states so that the existing and the being-developed analysis procedures can be properly applied. The class of systems under consideration includes general nonlinear systems with linear junction structures.

The second topic is the identification of relative degrees and zero dynamics. Relative degrees and zero dynamics are important features for the design of feedback control laws. For certain systems, the zero dynamics even directly determines the performance limits. Since the intrinsic zero dynamics can not be influenced by any feedback compensation, it is important to design the physical systems so that they possess desired zero dynamics. However, the calculation of the zero dynamics is usu-

ally complicated, especially if a form which is closely related to the physical system and suitable for design is required. A ZDIP procedure is proposed to derive the zero dynamics of physical systems from bond graph models. This method incorporates the definition of zero dynamics in the differential geometric approach and the causality manipulation in the bond graph representation. By doing so, the state equations of the zero dynamics can be easily obtained. The system elements which are responsible for the zero dynamics can be identified. In addition, if isolated subsystems which exhibit the zero dynamics exist, they can be found. Thus, the design of physical systems including the consideration of the zero dynamics become straightforward. The class of systems under consideration includes general nonlinear MIMO systems with linear junction structures, one port energy storage and dissipative elements.

With suitable modifications, the proposed approach can handle the systems with multi-port fields and the systems with modulated junctions (nonlinear junctions) or modulated elements in a similar manner. In the first case, part of the junction information is embedded in the field constitutive equations. Therefore, additional information from the fields is necessary for the search of the shortest causal paths. The differential geometric approach can be employed to provide such information. The same procedure can then be applied to identify the relative degrees and the zero dynamics. In the second case, if the shortest causal path only go through the energy bonds, the proposed procedure can be applied without any modification. However, if the shortest causal paths do go through the information bonds which are associated with the modulated elements, the causality manipulation may need to be modified so that the variable dependency is properly represented by the zero dynamics model. In both cases, the modified procedures should be consistent with the proposed ZDIP procedure.

The purpose of the third topic is to build the direct relations between the com-

ponent characteristics and the system eigenvalues. It is known that the symbolic solutions for the eigenvalues of high order systems are not available. Even if the exact solutions exist, they may be too complicated, and therefore do not point out useful design directions. In this thesis, several decomposition procedures are proposed to identify the physical components which contribute most to certain group of eigenvalues. By using the available matrix theories, the bounds of each eigenvalue group can be represented in terms of the component characteristics. These bounds will then facilitate the design of physical systems so that they have the eigenvalues roughly at the desired locations. The class of systems under consideration includes linear systems with decomposable eigenvalue groups. For general linear systems with heavily coupled eigenvalue groups, the proposed decomposition procedures may not provide a satisfactory solution. In this case, by appropriate partitions, useful eigenvalue bounds may be obtained by the proposed pre-conditioning procedure. Since there is no systematic partitioning procedure to guarantee this result, further study is recommended in this direction.

From the results of this research, it is shown that the analysis and design of dynamic systems can be conducted in a systematic way by studying the system configurations. The proposed procedures are ready to be coded and be included into a computer-aided design package.

Proofs concerning the explicit fields

A.1 Independent state variables contributed by explicit fields

Proof of Lemma 2.1:

1. Consider the constitutive equations of the field in Eqn.(2.7).

$$\begin{pmatrix} \mathbf{e}_{out} \\ \mathbf{q}_{out} \end{pmatrix} = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{bmatrix} \begin{pmatrix} \mathbf{q}_{in} \\ \mathbf{e}_{in} \end{pmatrix}$$

Suppose the imposed flows to the \mathbf{q}_{in} ports are \mathbf{f}_{in} . The state equations associated with this field would be

$$\dot{\mathbf{q}}_{in} = \mathbf{f}_{in} \tag{A.1}$$

Once \mathbf{q}_{in} are obtained by integration, the output variables \mathbf{e}_{out} , \mathbf{q}_{out} can be determined by Eqn.(2.7). However, if \mathbf{C}_{11} is not a full rank matrix, then the output variables \mathbf{e}_{out} can not be determined independently by the inputs \mathbf{f}_{in} . Namely, the inputs \mathbf{f}_{in} can drive \mathbf{q}_{in} freely, but not \mathbf{e}_{out} . This is because the following equation do not exist if \mathbf{C}_{11} is not invertible.

$$\mathbf{q}_{in} = \mathbf{C}_{11}^{-1} \mathbf{e}_{out} + \mathbf{C}_{11}^{-1} \mathbf{C}_{12} \mathbf{e}_{in} \tag{A.2}$$

Furthermore, by matrix theories [16], the number of variables in \mathbf{e}_{out} which can be determined independently by the inputs \mathbf{f}_{in} is equal to the rank of matrix \mathbf{C}_{11} . Therefore, the number of independent states this field can contribute is determined by the rank of the submatrix \mathbf{C}_{11} . □

2. Suppose that Eqn.(2.7) is written as the following by making $\mathbf{e}_{out} = \mathbf{e}_1$, $\mathbf{e}_{in} = \mathbf{e}_2$, $\mathbf{q}_{in} = \mathbf{q}_1$, $\mathbf{q}_{out} = \mathbf{e}_2$.

$$\begin{pmatrix} \mathbf{e}_1 \\ \mathbf{q}_2 \end{pmatrix} = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{bmatrix} \begin{pmatrix} \mathbf{q}_1 \\ \mathbf{e}_2 \end{pmatrix}$$

If the submatrices \mathbf{C}_{11} and \mathbf{C}_{22} are full rank, Eqn.(2.7) can be transformed into the following form. This is equivalent to assigning all integral causalities to the field.

$$\begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \end{pmatrix} = \begin{bmatrix} \mathbf{C}_{11} - \mathbf{C}_{12}\mathbf{C}_{22}^{-1}\mathbf{C}_{21} & \mathbf{C}_{12}\mathbf{C}_{22}^{-1} \\ -\mathbf{C}_{22}^{-1}\mathbf{C}_{21} & \mathbf{C}_{22}^{-1} \end{bmatrix} \begin{pmatrix} \mathbf{q}_1 \\ \mathbf{q}_2 \end{pmatrix} = \mathbf{C}_{int} \begin{pmatrix} \mathbf{q}_1 \\ \mathbf{q}_2 \end{pmatrix} \quad (\text{A.3})$$

Also, in the same manner, Eqn.(2.7) can be transformed into the following form. This is equivalent to assigning all derivative causalities to the field.

$$\begin{pmatrix} \mathbf{q}_1 \\ \mathbf{q}_2 \end{pmatrix} = \begin{bmatrix} \mathbf{C}_{11}^{-1} & -\mathbf{C}_{11}^{-1}\mathbf{C}_{12} \\ \mathbf{C}_{21}\mathbf{C}_{11}^{-1} & \mathbf{C}_{22} - \mathbf{C}_{21}\mathbf{C}_{11}^{-1}\mathbf{C}_{12} \end{bmatrix} \begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \end{pmatrix} = \mathbf{C}_{der} \begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \end{pmatrix} \quad (\text{A.4})$$

By the above two equations, the following relations can be obtained.

$$\mathbf{C}_{der} = \mathbf{C}_{int}^{-1} \quad (\text{A.5})$$

$$\mathbf{C}_{int} = \mathbf{C}_{der}^{-1} \quad (\text{A.6})$$

Therefore, the matrix \mathbf{C}_{int} and \mathbf{C}_{der} are full rank.

Suppose that a certain causality assignment to this field is desired, the associated constitutive equations can be derived by partitioning the matrix \mathbf{C}_{int} (or \mathbf{C}_{der} in the same manner) as follows.

$$\begin{pmatrix} \mathbf{e}'_1 \\ \mathbf{e}'_2 \end{pmatrix} = \begin{bmatrix} \mathbf{C}'_{11} & \mathbf{C}'_{12} \\ \mathbf{C}'_{21} & \mathbf{C}'_{22} \end{bmatrix} \begin{pmatrix} \mathbf{q}'_1 \\ \mathbf{q}'_2 \end{pmatrix} \quad (\text{A.7})$$

The new constitutive equations are

$$\begin{pmatrix} \mathbf{e}'_1 \\ \mathbf{q}'_2 \end{pmatrix} = \begin{bmatrix} \mathbf{C}'_{11} - \mathbf{C}'_{12}\mathbf{C}'_{22}{}^{-1}\mathbf{C}'_{21} & \mathbf{C}'_{12}\mathbf{C}'_{22}{}^{-1} \\ -\mathbf{C}'_{22}{}^{-1}\mathbf{C}'_{21} & \mathbf{C}'_{22}{}^{-1} \end{bmatrix} \begin{pmatrix} \mathbf{q}'_1 \\ \mathbf{e}'_2 \end{pmatrix} \quad (\text{A.8})$$

The field can not accept this causality assignment if and only if the submatrix \mathbf{C}'_{22} is not invertible. However, if the submatrix \mathbf{C}'_{22} is one of \mathbf{C}_{int} 's diagonal sub-blocks.

If it is not invertible, the matrix \mathbf{C}_{int} is not full rank. This contradicts with the fact found in the above. Therefore, this field can accept any combination of causality assignment. Also, combining with Lemma 2.1.1, it is certain that the number of the independent states contributed by this field is indicated by the number of ports with integral causalities. \square

Proof of Lemma 2.2:

Similar to the discussion in the proof of Lemma 2.2.2, if we would like to reverse the causality of the ports associated with variables \mathbf{e}'_2 in Eqn.(A.7), then the inverse of the submatrix \mathbf{C}'_{22} must exist as shown in Eqn.(A.8). For nonlinear fields, the inverse of the associated Jacobian submatrix must exist. If \mathbf{C}'_{22} is not full rank, however, we can repartition matrix \mathbf{C}_{int} and get a full rank submatrix \mathbf{C}_{22}'' associated with part of the variables in \mathbf{e}'_2 . By matrix theories, the rank of \mathbf{C}_{22}'' is equal to the rank of \mathbf{C}'_{22} . The number of variables in \mathbf{e}'_2 are not included in the partition is the number of ports whose causality can be reversed. Therefore, the number of these ports indicates the rank deficiency of the submatrix \mathbf{C}'_{22} . \square

Proof of Proposition 2.1:

1. This statement is self-proven by the results of Lemma 2.1.1 and Lemma 2.2.1. \square
2. This statement is self-proven by the results of Lemma 2.1.2 and Lemma 2.2.1 \square
3. This statement is self-proven by the results of Lemma 2.1.1, 2.2.1 and the definition of type 2 excess states. \square

A.2 The coupling in explicit fields

As discussed in section 2.3, for general systems, the submatrices \mathbf{C}_{11} and \mathbf{C}_{22} are not necessarily strictly positive-definite, the causality assignment which can be accepted by the field might be constrained. In this case, the independent states which this field can contribute might be different from what the integral causalities indicate. This

is caused by the dependency of the bond variables associated with the field ports. Lemma 2.1 and lemma 2.2 are self-proven when deriving the state equations of a system and the constitutive equations of a field according to the causality. The following procedure further illustrates that such constraints in a field can be "released" into the junction structure. The result of this procedure is a new explicit field which can accept all kind of causality assignment (at least mathematically). All the constraints in the original field will be represented by augmented junction structures.

This procedure is demonstrated by the following example. Consider a C field with the constitutive relation as follows.

$$\begin{pmatrix} e_1 \\ e_2 \\ e_3 \\ q_4 \\ q_5 \end{pmatrix} = \left[\begin{array}{ccc|cc} d_{11} & d_{12} & d_{13} & a_{11} & a_{11} \\ d_{12} & d_{22} & d_{23} & a_{21} & a_{22} \\ d_{13} & d_{23} & d_{23} & a_{31} & a_{32} \\ \hline b_{11} & b_{12} & b_{13} & d_{44} & d_{45} \\ b_{21} & b_{22} & b_{23} & d_{54} & d_{55} \end{array} \right] \begin{pmatrix} q_1 \\ q_2 \\ q_3 \\ e_4 \\ e_5 \end{pmatrix} \quad (\text{A.9})$$

Suppose the submatrix $\begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{12} & d_{22} & d_{23} \\ d_{13} & d_{23} & d_{23} \end{bmatrix}$ is not full rank, i.e. the columns are dependent. For example, if the rank is 2, there are 2 and only 2 independent columns in this matrix. Without loss of generality, suppose the first and the second columns are independent, the third one can be represented as

$$\begin{pmatrix} d_{31} \\ d_{32} \\ d_{33} \end{pmatrix} = \alpha \begin{pmatrix} d_{11} \\ d_{12} \\ d_{13} \end{pmatrix} + \beta \begin{pmatrix} d_{21} \\ d_{22} \\ d_{23} \end{pmatrix} \quad (\text{A.10})$$

, where α and β are real numbers and not both zero at the same time. The constitutive relation can be written as

$$\begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} = \begin{pmatrix} d_{11} \\ d_{12} \\ d_{13} \end{pmatrix} (q_1 + \alpha q_3) + \begin{pmatrix} d_{21} \\ d_{22} \\ d_{23} \end{pmatrix} (q_2 + \beta q_3) + \begin{pmatrix} a_{11} \\ a_{22} \\ a_{33} \end{pmatrix} e_4 + \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix} e_5 \quad (\text{A.11})$$

Note that the submatrix $\begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{12} & d_{22} & d_{23} \\ d_{13} & d_{23} & d_{23} \end{bmatrix}$ is symmetric. Thus the following equations can be obtained.

$$e_3 = \alpha(e_1 - a_{11}e_4 - a_{12}e_5) + \beta(e_2 - a_{21}e_4 - a_{22}e_5) + a_{31}e_4 + a_{32}e_5 \quad (\text{A.12})$$

$$= \alpha e_1 + \beta e_2 + (-\alpha a_{11} - \beta a_{21} + a_{31})e_4 + (-\alpha a_{12} - \beta a_{22} + a_{32})e_5 \quad (\text{A.13})$$

$$= \alpha e_1 + \beta e_2 + \kappa_1 e_4 + \kappa_2 e_5 \quad (\text{A.14})$$

If the field is represented as Figure A.1, the explicit field equations become

$$\begin{pmatrix} e_1 \\ e_2 \\ q'_4 \\ q'_5 \end{pmatrix} = \left[\begin{array}{cc|cc} d_{11} & d_{12} & a_{11} & a_{11} \\ d_{12} & d_{22} & a_{21} & a_{22} \\ \hline b_{11} & b_{12} & d_{44} & d_{45} \\ b_{21} & b_{22} & d_{54} & d_{55} \end{array} \right] \begin{pmatrix} q_1 + \alpha q_3 \\ q_2 + \beta q_3 \\ e_4 \\ e_5 \end{pmatrix} \quad (\text{A.15})$$

By the above derivation, it is shown that the new constitutive equations together with the newly added junction structure are equivalent to the original explicit field. Therefore, the constraints are successfully represented by the junction structure. This is possible mainly because of the property $\mathbf{C}_{12} = -\mathbf{C}_{21}^T$ (shown in section 2.3). Since now the diagonal submatrices of the resultant explicit field are strictly positive-definite, the field will be able to accept any causality assignment.

An example

Consider a system from [25] as shown in Figure A.2-(a). If the implicit field form in Figure A.2-(b) are represented by an explicit field. The constitutive relation will be

$$\begin{pmatrix} f_5 \\ f_7 \\ \hline p_6 \\ p_8 \end{pmatrix} = \left[\begin{array}{cc|cc} \mathbf{I}_{11} & & \mathbf{I}_{12} & \\ \hline & & & \mathbf{I}_{22} \end{array} \right] \begin{pmatrix} p_5 \\ p_7 \\ \hline f_6 \\ f_8 \end{pmatrix}$$

$$= \left[\begin{array}{cc|cc} 0 & 0 & T_1 & T_1 \\ 0 & 0 & T_2 & -T_2 \\ \hline -T_1 & -T_2 & \frac{T_1^2}{I_1} + \frac{T_2^2}{I_2} + \frac{1}{I_3} & \frac{T_1^2}{I_1} - \frac{T_2^2}{I_2} \\ -T_1 & T_2 & \frac{T_1^2}{I_1} - \frac{T_2^2}{I_2} & \frac{T_1^2}{I_1} + \frac{T_2^2}{I_2} + \frac{1}{I_4} \end{array} \right] \begin{pmatrix} p_5 \\ p_7 \\ \hline f_6 \\ f_8 \end{pmatrix}$$

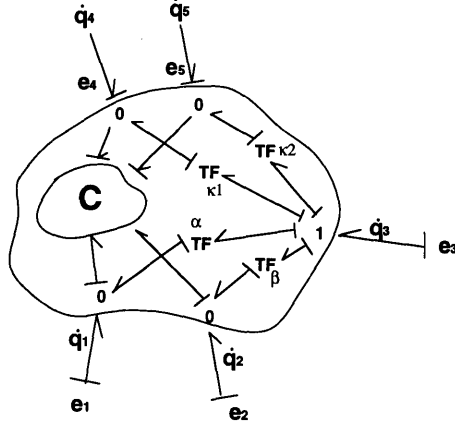


Figure A.1: A new representation of explicit field with no constraints inside.

Note that \mathbf{I}_{11} being $\mathbf{0}$ indicates that this field contributes no independent states. This is consistent with the prediction from the causalities in Figure A.2-(a) but not from the causalities in Figure A.2-(b). Obviously, the explicit field contains hidden constraints. Using the procedure described in the previous section, the constraints in this field can use represented by transformers and zero junctions as shown in Figure A.3-(c). The new I field now becomes

$$\begin{pmatrix} p'_6 \\ p'_8 \end{pmatrix} = \begin{bmatrix} \frac{T_1^2}{I_1} + \frac{T_2^2}{I_2} + \frac{1}{I_3} & \frac{T_1^2}{I_1} - \frac{T_2^2}{I_2} \\ \frac{T_1^2}{I_1} - \frac{T_2^2}{I_2} & \frac{T_1^2}{I_1} + \frac{T_2^2}{I_2} + \frac{1}{I_4} \end{bmatrix} \begin{pmatrix} f_6 \\ f_8 \end{pmatrix} \quad (\text{A.16})$$

This is a positive definite constitutive relation which can accept any combination of causality assignments. The causalities now correctly indicate the number of the independent states, which is 0 as shown before. Finally the graph can be simplified as shown in Figure A.3-(d).

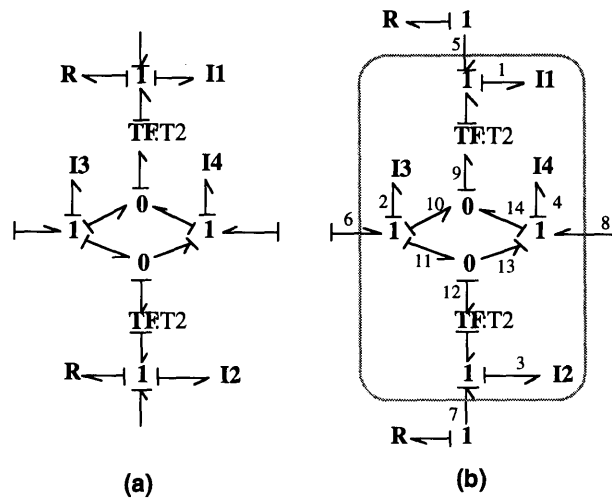


Figure A.2: (a) The original system. (b) An implicit field form.

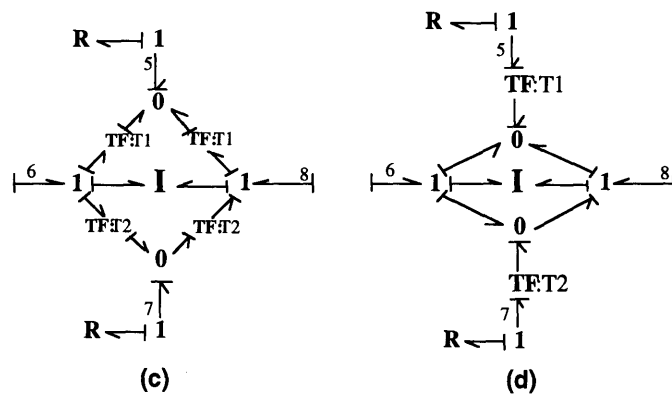


Figure A.3: (c) The system after releasing the constraints. (d) The final form.

Appendix B

Proofs concerning the relative degrees and the zero dynamics

B.1 Relative degrees

In [30, 47], it has been proved that the structural relative degree is the number of integral causality on the shortest causal path less the number of derivative causality on the same path. Thus, in what follows, proposition 3.1 will be proved and proposition 3.2 would be self-proved by the application of proposition 3.1 and the results in [30, 47].

Proposition 3.1 will be proved by the use of the following lemmas. For simplicity, the following statements consider only the equivalent bond graph models where elements have been reflected to the same energy domain.

Lemma B.1.1: In an SISO bond graph model where elements have been reflected to the same energy domain, if a causal path which connects the input to an output variable (a state variable) contains an energy storage element with derivative causality, there must be an energy storage element of the same type with integral causality on the same path and directly causally related¹ to this energy storage element.

Proof: According to the definition, the causal path can be represented by a sequence of bond variables, for example,

$$\text{input} \dots f_{in} - f_j - e_j - e_{j+1} - f_{j+1} - f_{out} \dots \text{output}$$

¹Two elements are said to be directly causally related if there is a causal path between these two elements without going through elements other than the junctions.

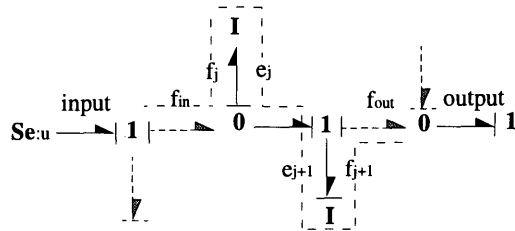


Figure B.1: A causal path which contains a derivative causality.

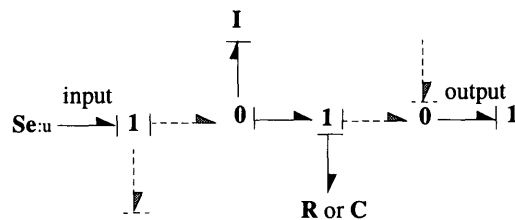


Figure B.2: An alternative causality assignment.

,where e 's represents efforts and f 's represents flows as shown in Figure B.1. For generality, the dashed bonds and causalities indicate the possibility of other valid structures. The variable type changes only when the path passes through the one port R , C , or I element. Suppose that a causal path which connects the input to an output variable (a state variable) contains an I element with derivative causality such as the one in the model of Figure B.1, the input variable to this element would be a flow variable and the output variable would be an effort variable. For the sequence to go on, the input variable to the next one port element on the path must be an effort variable. If this next element was an R element or a C element, according to the Sequential Causality Assignment Procedure, the I element would have been assigned an integral causality as shown in Figure B.2. Thus this next element on the causal path must be an I element with integral causality. Note that same arguments apply to C elements.

Lemma B.1.2: If a causal path which connects the input to an output variable (a state variable) contains an energy storage element with derivative causality, an alter-

native causality assignment exists such that the new causal path which connects the same input and output does not pass through this element with derivative causality.

Proof: Suppose that this element with derivative causality is an I element, the causal path would be

$$\text{input} \dots - f_{in} - f_j - e_j - e_{j+1} - f_{j+1} - f_{out} \dots \text{-output}$$

,where f_j, e_j are the input and output variable of the I element with derivative causality, e_{j+1}, f_{j+1} are the input and output variable of the I element with integral causality. Since these two I elements are directly causally connected from lemma B.1.1, there exists an alternative causality assignment with the causality of both I elements reversed as shown in Figure B.3. By reversing the causality of this two I elements at the same time, the input and output variables of the I elements switches and the sequence becomes

$$\text{input} \dots - f_{in} - f_{out} \dots \text{-output}$$

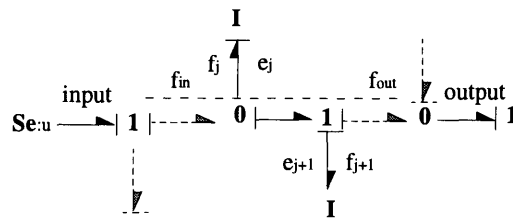


Figure B.3: An alternative causality assignment.

Thus this causal path does not pass through these two I elements. Note that same arguments apply to C elements.

Proof of proposition 3.1:

By lemma B.1.1 and lemma B.1.2, if a shortest causal path contains a derivative causality, an alternative causality assignment exists so that the new shortest causal path becomes a simple causal path. Thus the statement in proposition 3.1 is true. \square

B.2 Supplements to the ZDIP procedure

Suppose the output y is one of the states (bond variable of an energy storage element). From the junction structure, \dot{y} can be represented as a function of the output variables (\mathbf{w}_1) of the energy storage elements which are semi-directly causally related² to the output y . Thus, \ddot{y} can be represented as a function of $\dot{\mathbf{w}}_1$. The variables in the vector $\dot{\mathbf{w}}_1$ can then be written as functions of the output variables (\mathbf{w}_2) of the energy storage elements which are semi-directly causally related to the elements involved in vector \mathbf{w}_1 . By this derivation, if the differentiation continues, one and only one new bond variable of the energy storage element on the shortest causal path appears in (\mathbf{w}_i) after each differentiation. It can be concluded that $y^{(k)}$ can be represented as a function of \mathbf{w}_k , which contains the *output* variable of the k^{th} energy storage element on the shortest causal path from the output to the input. Equivalently, $y^{(k)}$ can be represented as a function of $\dot{\mathbf{w}}_{k-1}$, which contains the *input* variable of the k^{th} energy storage element on the shortest causal path from the output to the input. Therefore, if the constraints $y = 0, \dot{y} = 0, \dots, y^{(r-1)} = 0$ are imposed to the bond graph, the corresponding graph can be represented by making the *input* variables of the energy storage elements on the shortest causal path dependent on other variables in $\dot{\mathbf{w}}_{k-1}$. That is, the input variables, which are originally determined by the system, now become output variables, which are determined by the constraints.

B.3 The vector relative degrees

Proof of proposition 4.1:

If any one of the inputs is in the condition as described in proposition 4.1, i.e. all outputs can be connected to other inputs by shorter causal paths than to this one, the

²Two energy storage elements are said to be semi-directly causally related if there is a causal path between these two elements without going through any other energy storage element.

corresponding column to this input in the decoupling matrix in section 4.2 will contain only zero elements. For example, if input j is such an input, then $L_{gj}L_{\mathbf{f}}^{r_i-1}h_i(\mathbf{x}) = 0$ for all $1 \leq i \leq m$. This is because the smallest k_i for $L_{gj}L_{\mathbf{f}}^{k_i}h_i(\mathbf{x})$ to be non-zero for each output i is always larger than $r_i - 1$. Thus the decoupling matrix is singular at any operating point. The vector relative degree of such a system can not be defined. \square

Proof of proposition 4.2:

This proposition will be proved by the following description and the lemmas listed below.

Assume that the shortest causal paths of input-output pairs u_1, y_1 and u_2, y_2 partially overlap. The energy storage elements (all with integral causalities) on the non-overlapping part of the shortest causal path beginning from the input u_1 side are denoted as $A_1 \dots A_p$. The energy storage elements on the non-overlapping part of the shortest causal path beginning from the input u_2 side are denoted as $B_1 \dots B_q$. If the differentiations of output y_1 are taken consecutively, to some order k , the bond variables of the energy storage elements A_p and B_q will explicitly appear. Namely, $y_1^{(k)}$ is the function of the bond variables of A_p and B_q . If the differentiation continues, then the bond variables of $A_{p-1}, B_{q-1}; A_{p-2}, B_{q-2} \dots$ appear consecutively. Finally, u_1 appears explicitly first if $p < q$ and u_2 appears explicitly first if $q < p$. Similarly, if the differentiations of y_2 are taken, to some order ℓ , the bond variables of the energy storage elements A_p and B_q will explicitly appear. Thus the same sequence $A_{p-1}, B_{q-1}; A_{p-2}, B_{q-2} \dots$ results. Similarly, u_1 appears explicitly first if $p < q$ and u_2 appears explicitly first if $q < p$.

Lemma B.3.1: If a system has partially overlapped shortest causal paths as described above, u_1 and u_2 will appear at the same order of differentiation of $y_i, i = 1, 2$, i.e. $p = q$.

Proof: Since u_1, y_1 and u_2, y_2 are taken as input-output pairs on shortest causal paths, the system configuration in proposition 4.1 has been ruled out (and one would use dynamic extension procedure if necessary). Suppose that $p \neq q$, then the differentiations of y_1 and y_2 both would have reached only u_1 or only u_2 first. This contradicts the previous statement. Thus $p = q$ is guaranteed. Note that one of u_1 and u_2 might be the nominal input from dynamic extension.

Lemma B.3.2: If a system has partially overlapped shortest causal paths as described above, the row vectors $\mathbf{D}_1 = [L_{g1}L_{\mathbf{f}}^{r_1-1}h_1(\mathbf{x}), L_{g1}L_{\mathbf{f}}^{r_2-1}h_2(\mathbf{x})]$ and $\mathbf{D}_2 = [L_{g2}L_{\mathbf{f}}^{r_1-1}h_1(\mathbf{x}), L_{g2}L_{\mathbf{f}}^{r_2-1}h_2(\mathbf{x})]$ will be structurally dependent.

Proof: From the previous lemma, $p = q$ is guaranteed. Thus all elements in the vectors $\mathbf{D}_1(\mathbf{x})$ and $\mathbf{D}_2(\mathbf{x})$ are non-zero. However, since the inputs u_1 and u_2 appear by differentiating the bond variables of the energy storage elements A_p and B_q for both outputs, the two row vectors are related by

$$\mathbf{D}_1(\mathbf{x}) = \alpha(\mathbf{x})\mathbf{D}_2(\mathbf{x}) \quad (\text{B.1})$$

, where $\alpha(\mathbf{x})$ is a scalar function of \mathbf{x} . Namely, $\mathbf{D}_1(\mathbf{x})$ and $\mathbf{D}_2(\mathbf{x})$ are structurally dependent.

Lemma B.3.3: If a system has partially overlapped shortest causal paths as described above and there are no other ways to avoid this overlapping, the decoupling matrix is singular.

Proof: Suppose that the considered system has more than two inputs and outputs, the singularity of the decoupling matrix can not be determined only by the row vectors in the previous lemma. However, if any of the column vectors $[L_{gj}L_{\mathbf{f}}^{r_1-1}h_1(\mathbf{x}) \ L_{gj}L_{\mathbf{f}}^{r_2-1}h_2(\mathbf{x})]^T, j > 2$ contains non-zero elements and are independent to the column vectors $[L_{gj}L_{\mathbf{f}}^{r_1-1}h_1(\mathbf{x}) \ L_{gj}L_{\mathbf{f}}^{r_2-1}h_2(\mathbf{x})]^T, j = 1, 2$, it would be possible to select an alternative shortest causal path for y_1 or y_2 to the corresponding u_j so that the shortest

causal paths for y_1 and y_2 are non-overlapping. This contradicts the statement of this lemma. Thus the column vectors $[L_{gj}L_{\mathbf{f}}^{r_1-1}h_1(\mathbf{x}) \ L_{gj}L_{\mathbf{f}}^{r_2-1}h_2(\mathbf{x})]^T$, $j > 2$ are either zero vectors or are dependent to $[L_{gj}L_{\mathbf{f}}^{r_1-1}h_1(\mathbf{x}) \ L_{gj}L_{\mathbf{f}}^{r_2-1}h_2(\mathbf{x})]^T$, $j = 1, 2$. This indicates that the corresponding row vectors for y_1 and y_2 in the decoupling matrix are dependent. Therefore, the decoupling matrix is singular.

By lemma B.3.1, B.3.2, and B.3.3, the decoupling matrix will be singular if a system has partially overlapped shortest causal paths as described above and there are no other ways to avoid this sharing. Thus the statement of this proposition is true. □

Appendix C

A maple procedure

To verify the results of the proposed analysis procedures in this thesis, the state equations of all the bond graph models are derived by a MAPLE¹ procedure. The use of this procedure is shown by an example in Figure C.1. This is a simplified version of the arm prosthesis model discussed in section 3.4. The corresponding command for the derivation of the state equations in MAPLE is shown below.

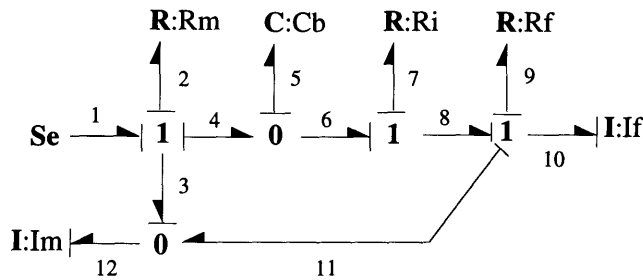


Figure C.1: The bond graph model of an arm prosthesis.

```
pp:=bdlin(5, [Se,R,C,R,R,I,I], [1,2,5,7,9,10,12],
  [-3,1,-2,-4], 1, [-5,4,-6], 0, [-8,6,-7], 1, [-10,8,-9,-11], 1, [3,11,-12], 0);
```

First of all, the bonds in the model are numbered by the user. According to these numbers, the user will input the bond graph structures by a set of data strings. The first argument is the number of junctions contained in this model. The second argument is a list which represents the types of the elements in the model. The third

¹MAPLE is a software package for mathematical symbolic derivations by Waterloo Maple Company .

argument is a list of numbers which represent the corresponding bonds connected to the elements in the previous list. The rest of the arguments describe the junctions and the bonds connected to them. For example, the next two arguments: $[-3, 1, -2, -4]$, 1 show that bonds 1, 2, 3 and 4 are connected to a one junction where the power directions of bonds 2, 3 and 4 are pointing out of the junction and the power direction of bond 1 is pointing into the junction. The first number in the list also shows the bond which dominates the causality of this junction. For this one junction, bond 3 imposes the flow. Therefore, the causalities of other bonds are determined. The other junctions are described in the same manner by consecutive arguments. This procedure is designed for the derivation of linear state equations. The resultant A and B matrices of this model are shown in Figure C.2. The source codes of this procedure are listed in the following pages. For nonlinear systems, a similar procedure in MAPLE is used to verify the analysis results.

$$\begin{bmatrix} 0 & -2 L10 & L12 \\ \frac{2}{C5} & (-R2 - R7 - R9) L10 & R2 L12 \\ 1 & R2 L10 & -R2 L12 \\ -\frac{1}{C5} & & \end{bmatrix} \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$$

Figure C.2: The results from the MAPLE procedure.

```

#####
#
#   This MAPLE procedure derives the state equations automatically
#   from a set of data strings which describe the bond graph
#   structures.
#
#   bdlin(jn,eletype,outbond,bonds1,jn1,bonds2,jn2.....);
#   jn : number of junctions
#   eletype: the element list in the bond graph
#   outbond: the bonds which are connected to the eletype elements
#   bonds1: the bonds which are connected to the first junction
#           + power goes into the junction
#           - power goes out of the junction
#           the first bond dominates the causality
#   jn1: the junction type of the first junction (0 or 1)
#   .... continue the bonds# and jn# strings
#
#
#   Version 1.30 Jan,1995 by Shih-Ying Huang, Copyright Reserved
#
#####

bdlin:=proc()
local jn,eletype,outbond,obnumber,etype,jun,bnumber,bind,eln,junn,
inv,efin,prej,tempj,i,efout,perm,swi,pirt,j,ptest,outv,prejn,inbnumber,
j11,j12,j21,j22,II,junction,efins,efouts,efoutse,Cn,In,DCn,DIn,Rn,Gn,Sn,invv,
juneqn,CImatrix,DCImatrix,RGmatrix,S11,S12,S13,S14,S21,S24,S31,S32,S33,S34,Iden,
A,B,TPKsd,Ksd,EU;

# read in the input strings

jn:=args[1];
eletype:=args[2]:
outbond:=args[3]:
obnumber:=nops(outbond):
etype:=array(1..jn):
jun:=array(1..jn):
bnumber:=array(1..jn)
bind:=0:

# assembling the junctions and the state vectors
for eln from 1 to jn do
  jun[eln]:=args[3+(eln-1)*2+1];
  etype[eln]:=args[3+(eln-1)*2+2]:
  bnumber[eln]:=nops(jun[eln]):

```

```

junn:=junn[eln]:

if eln=1
then inv:=matrix(1,bnumber[eln],0):
     efin:=matrix(1,bnumber[eln],0):
else inv:=extend(inv,0,bnumber[eln],0):
     efin:=extend(efin,0,bnumber[eln],0):
fi:

if eln=1
then prej:=matrix(bnumber[eln],bnumber[eln],0):
else prej:=extend(prej,bnumber[eln],bnumber[eln],0):
fi:

if 1=1 then
tempj:=matrix(bnumber[eln],bnumber[eln],0):
fi:

if etype[eln]=0 or etype[eln]=1
then for i from 2 to bnumber[eln] do tempj[i,1]:=1 od:
     for i from 2 to bnumber[eln] do
         tempj[1,i]:=csgn(junn[1])*(-1)*csgn(junn[i]) od:
elif etype[eln]=g
then tempj[1,2]:=cat(g,convert(eln,string)):
     tempj[2,1]:=cat(g,convert(eln,string)):
elif etype[eln]=t
then tempj[1,2]:=cat(t,convert(eln,string)):
     tempj[2,1]:=cat(t,convert(eln,string)):
fi:

if eln<>1 then
     bind:=bind+bnumber[eln-1]:
fi:

if 1=1 then
     copyinto(tempj,prej,bind+1,bind+1):
fi:

for i to bnumber[eln] do inv[1,i+bind]:=abs(junn[i]) od:

if etype[eln]=0
then efin[1,bind+1]:=1:
elif etype[eln]=1 then for i from 2 to bnumber[eln] do efin[1,i+bind]:=1 od:
elif etype[eln]=g then for i from 1 to bnumber[eln] do
     if junn[1]>0 then efin[1,i+bind]:=1

```

```

        else efin[1,i+bind]:=0 fi: od:
    elif etype[eln]=t then
        if junn[1]>0 then efin[1,(1+bind)]:=1:
            efin[1,(2+bind)]:=0:
            else efin[1,(1+bind)]:=0:
                efin[1,(2+bind)]:=1: fi:
        fi:
    od:

bind:=bind+bnumber[jn]:
efout:=matrix(1,bind,0):
for i to bind do if efin[1,i]=1 then efout[1,i]:=0
    else efout[1,i]:=1 fi: od:
perm:=matrix(bind,bind,0):
swi:=array(1..bind):
pirt:=0:

for i to obnumber do
    for j to bind do
        if inv[1,j]=outbond[i]
            then pirt:=pirt+1: swi[pirt]:=j:
        fi:
    od:
od:

pirt:=0:
ptest:=0:

for i to bind do
    for j to obnumber do
        if inv[1,i]=outbond[j]
            then ptest:=1:
        fi:
    od:
    if ptest<>1 then
        pirt:=pirt+1:
        swi[pirt+obnumber]:=i:
    fi:
    ptest:=0:
od:

for i to bind do perm[i,swi[i]]:=1 od:
inv:=transpose(multiply(perm,transpose(inv))):
efin:=transpose(multiply(perm,transpose(efin))):
efout:=transpose(multiply(perm,transpose(efout))):

```



```

outv:=inv:
prejn:=multiply(perm,prej,inverse(perm)):
inbnumber:=bind-obnumber:

j11:=submatrix(prejn,1..obnumber,1..obnumber):
j12:=submatrix(prejn,1..obnumber,(obnumber+1)..bind):
j21:=submatrix(prejn,(obnumber+1)..bind,1..obnumber):
j22:=submatrix(prejn,(obnumber+1)..bind,(obnumber+1)..bind):

perm:=matrix(inbnumber,inbnumber,0):
swi:=array(1..inbnumber):

for i from (obnumber+1) to bind do
  for j from (obnumber+1) to bind do
    if efout[1,j]=efin[1,i] and inv[1,i]=outv[1,j] then
      swi[i-obnumber]:=j-obnumber:
    fi:
  od:
od:

for i from 1 to inbnumber do perm[i,swi[i]]:=1 od:
II:=matrix(inbnumber,inbnumber,0):
for i to inbnumber do II[i,i]:=1 od:

# assembling the junction equations
junction:=
matadd(j11,multiply(j12,perm,
  inverse(matadd(II,scalarmul(multiply(j22,perm),-1))),j21)):

efins:=array(1..obnumber):
efouts:=array(1..obnumber):

for i to obnumber do
  if efin[1,i]=1 then efins[i]:=cat(e,convert(inv[1,i],string)):
  else efins[i]:=cat(f,convert(inv[1,i],string)) fi:
  if efout[1,i]=1 then efouts[i]:=cat(e,convert(inv[1,i],string)):
  else efouts[i]:=cat(f,convert(inv[1,i],string)) fi:
od:

# Permute the junction equations for the derivation of state equations
efoutse:=multiply(junction,efins):
perm:=matrix(obnumber,obnumber,0):
swi:=array(1..obnumber):
pirt:=0:

```

```

for i to obnumber do
  if eletype[i]=C and efin[1,i]=1 then pirt:=pirt+1: swi[pirt]:=i: fi:
od:

Cn:=pirt:
for i to obnumber do
  if eletype[i]=I and efin[1,i]=0 then pirt:=pirt+1: swi[pirt]:=i: fi:
od:

In:=pirt-Cn:

for i to obnumber do
  if eletype[i]=C and efin[1,i]=0 then pirt:=pirt+1: swi[pirt]:=i: fi:
od:

DCn:=pirt-Cn-In:

for i to obnumber do
  if eletype[i]=I and efin[1,i]=1 then pirt:=pirt+1: swi[pirt]:=i: fi:
od:

DIn:=pirt-DCn-Cn-In:

for i to obnumber do
  if eletype[i]=R and efin[1,i]=1 then pirt:=pirt+1: swi[pirt]:=i: fi:
od:

Rn:=pirt-DCn-DIn-Cn-In:

for i to obnumber do
  if eletype[i]=R and efin[1,i]=0 then pirt:=pirt+1: swi[pirt]:=i: fi:
od:

Gn:=pirt-DCn-DIn-Cn-In-Rn:

for i to obnumber do
  if eletype[i]=Sf or eletype[i]=Se then pirt:=pirt+1: swi[pirt]:=i: fi:
od:

Sn:=pirt-DCn-DIn-Rn-Cn-In-Gn:
for i from 1 to obnumber do perm[i,swi[i]]:=1 od:

# perform permutations
efins:=(multiply(perm,efins)):
efouts:=(multiply(perm,efouts)):

```

```

efoutse:=(multiply(perm,efoutse)):
invv:=submatrix(inv,1..1,1..obnumber):
invv:=transpose(multiply(perm,transpose(invv))):

# assembling the junction equations with permuted states
juneqn:=multiply(perm,junction,inverse(perm)):

# preparing the matrices describing the element characteristics
if (Cn+In)>=1 then
  CImatrix:=matrix((Cn+In),(Cn+In),0):
  for i to Cn do CImatrix[i,i]:=1/cat(C,convert(invv[1,i],string)) od:
  for i to In do CImatrix[i+Cn,i+Cn]:=cat(L,convert(invv[1,i+Cn],string))
od: fi:

if (DCn+DIn)>=1 then
  DCImatrix:=matrix((DCn+DIn),(DCn+DIn),0):
  for i to DCn do
    DCImatrix[i,i]:=cat(C,convert(invv[1,i+Cn+In],string))
  od:
  for i to DIn do
    DCImatrix[i+DCn,i+DCn]:=1/cat(L,convert(invv[1,i+DCn+Cn+In],string))
  od:
fi:

if (Rn+Gn)>=1 then RGmatrix:=matrix((Rn+Gn),(Rn+Gn),0):
  for i to Rn do RGmatrix[i,i]:=cat(R,convert(invv[1,i+Cn+In+DCn+DIn],string))
  od:
  for i to Gn
    do RGmatrix[i+Rn,i+Rn]:=1/cat(R,convert(invv[1,i+Cn+In+Rn+DCn+DIn],string)):
  od:
fi:

# assembling the submatrices for the state equation derivation
S11:=submatrix(juneqn,1..(Cn+In),1..(Cn+In)):
if (DCn+DIn)>=1 then
  S12:=submatrix(juneqn,1..(DCn+DIn),(Cn+In)+1..(Cn+In+DCn+DIn)):
fi:

if (Rn+Gn)>=1 then
  S13:=submatrix(juneqn,1..(Cn+In),Cn+In+DCn+DIn+1..(Cn+In+DCn+DIn+Rn+Gn)):
fi:

S14:=submatrix(juneqn,1..(Cn+In),(Cn+In+DCn+DIn+Rn+Gn)+1..(Cn+In+DCn+DIn+Rn+Gn)+Sn):
if (DCn+DIn)>=1 then
  S21:=submatrix(juneqn,(Cn+In)+1..(Cn+In+DCn+DIn),1..(Cn+In)):

```

```

S24:=submatrix(juneqn,(Cn+In)+1..(Cn+In+DCn+DIn),
              (Cn+In+DCn+DIn+Rn+Gn)+1..(Cn+In+DCn+DIn+Rn+Gn)+Sn):
fi:

if (Rn+Gn)>=1 then
  S31:=submatrix(juneqn,(Cn+In+DCn+DIn)+1..(Cn+In+Rn+Gn+DCn+DIn),1..(Cn+In)):
  if (DCn+DIn)>=1 then
    S32:=submatrix(juneqn,(Cn+In+DCn+DIn)+1..(Cn+In+Rn+Gn+DCn+DIn),
                  (Cn+In)+1..(Cn+In+DCn+DIn)):
    fi:
    S33:=submatrix(juneqn,(Cn+In+DCn+DIn)+1..(Cn+In+Rn+Gn+DCn+DIn),
                  Cn+In+DCn+DIn+1..(Cn+In+DCn+DIn+Rn+Gn)):
    S34:=submatrix(juneqn,(Cn+In+DCn+DIn)+1..(Cn+In+Rn+Gn+DCn+DIn),
                  (Cn+DCn+DIn+In+Rn+Gn)+1..(Cn+In+DCn+DIn+Rn+Gn)+Sn):
    fi:

  if (Rn+Gn)>=1 then
    Iden:=matrix((Rn+Gn),(Rn+Gn),0):
    for i to (Rn+Gn) do Iden[i,i]:=1; od:
  fi:

# assembling the A,B matrix
if (Rn+Gn)>=1 then
  A:=multiply(matadd(S11,multiply(S13,RGmatrix,
    inverse(matadd(Iden,scalarmul(multiply(S33,RGmatrix),-1))),S31)),CI matrix);
  B:=matadd(S14,multiply(S13,RGmatrix,
    inverse(matadd(Iden,scalarmul(multiply(S33,RGmatrix),-1))),S34));
else
  A:=multiply(S11,CI matrix);
  B:=(S14);
fi:

# assembling the A,B matrix if there are derivative causalities
if (DCn+DIn)>=1 then
  Iden:=matrix((DCn+DIn),(DCn+DIn),0):
  for i to (DCn+DIn) do Iden[i,i]:=1; od:
fi:

if (DCn+DIn)>=1 then
  TPKsd:=scalarmul(multiply(S12,DCI matrix),-1):
  Ksd:=multiply(S13,RGmatrix,
    inverse(matadd(Iden,scalarmul(multiply(S33,RGmatrix),-1))),S32,DCI matrix):
  Ksd:=matadd(TPKsd,Ksd):

A:=multiply(inverse(matadd(Iden,multiply(Ksd,S21,CI matrix))),A):

```

```
B:=multiply(inverse(matadd(Iden,multiply(Ksd,S21,CImatrix))),B):  
EU:=multiply(inverse(matadd(Iden,multiply(Ksd,S21,CImatrix))),Ksd,S24):  
RETURN(A,B,EU);  
  
end;
```

Bibliography

- [1] Abul-Haj, C., Hogan, N. "An Emulator System for Developing Improved Elbos-Prosthesis Designs", IEEE Transactions on Biomedical Engineering, Vol.BME-34, No.9. Sep. 1987.
- [2] Andry,Jr. A.N., Rosenberg, R.C. " On the Dimension of State Space for Physical Systems", Proceedings of JACC, 1983.
- [3] Barreto, J., Lefevre, J. " R-fields in the Solution of Implicit Equations", Journal of franklin Institute, 1985.
- [4] "Large Scale Matrix Problems", edited by Ake Bjorck, Robert J.Plemmons, and Hans Schneider, Elsevier North Holland Inc., New York, 1981.
- [5] Bos, A.M. " Implicit Solutions of Equations Derived from Mechanical Bond Graphs", Complex and Distributed Systems: Analysis, Simulation and Control, IMACS, 1986.
- [6] Breedveld, P.C., Rosenberg, R.C. and Zhou, T., "Bond Graph Bibliography", Special Issue Journal of the Franklin Institute on "Current Research in Bond Graph Modeling", Vol.328, No. 5/6, pp.1067-1109, 1991.
- [7] Breedveld, P.C. " Decomposition of Multiport Elements in a Revised Multibond Graph Notation", Journal of franklin Institute, Vol.318, No.2, pp. 77-89, 1984.
- [8] Breedveld, P.C. "Physical Systems Theory in terms of Bond Graphs", Thesis, Enschede, 1984.

- [9] Callen, H.B., "Thermodynamics and an Introduction to Thermostatistics", John Wiley & Sons, New York, 1985.
- [10] Doyle, J.C., Francis, B.A., Tannenbaum, A.R. "Feedback Control Theory", Maxwell Macmillan, 1992.
- [11] "Sparsity and Its Applications", edited by Davis J. Evans, Cambridge University Press, New York, 1985.
- [12] Filippo, M.J., Delgado, M., Brie, C. and Paynter, H.M. "A Survey of Bond Graphs: Theory, Application and Programs", Journal of Franklin Institute, Vol.328, No.5/6, pp.565-606, 1991.
- [13] Gawthrop, P.J., Smith, L. "Causal Augmentation of Bond Graphs with Algebraic Loops", Journal of Franklin Institute, Vol.329, No.2, pp. 291-303, 1992.
- [14] Hogan, N., "Modularity and Causality in Physical System Modelling", Journal of Dynamic Systems, Measurement, and control, Transactions of ASME, Vol. 109, pp. 384-391, Dec 1987.
- [15] Hogan, N., Fasse, E.D., "Conservation Principles and Bond Graph Junction Structures", Proc. ASME 1988 WAM, DSC-Vol. 8. pp. 9-14, 1988.
- [16] Horn, Roger A., Johnson, Charles R. "Matrix Analysis", Cambridge University Press, New York, 1985.
- [17] Huang, S.-Y. and Youcef-Toumi, K. "Explicit Fields and Their Applications to Structural Property Inspection of Physical Systems", Proceedings of IFAC, San Francisco, 1996.

- [18] Huang, S.-Y. and Youcef-Toumi, K. "Zero Dynamics of Nonlinear MIMO Systems from System Configurations - A Bond Graph Approach", Proceedings of IFAC, San Francisco, 1996.
- [19] Huang, S.-Y. and Youcef-Toumi, K. "Zero Dynamics of Physical Systems from Bond Graph Models - Part I : SISO Systems", Submitted to Journal of Dynamic Systems, Measurement, and Control, ASME.
- [20] Huang, S.-Y. and Youcef-Toumi, K. "Zero Dynamics of Physical Systems from Bond Graph Models - Part II : MIMO Systems", Submitted to Journal of Dynamic Systems, Measurement, and Control, ASME.
- [21] Isidori, Alberto, "Nonlinear Control Systems", 2nd edition, Springer-Verlag, 1989.
- [22] Karnopp, D., Margolis, D., Rosenberg, R. " System Dynamics: A Unified Approach", Second Edition, John Wiley & Sons, Inc., New York, 1990.
- [23] Karnopp, D. " On the Order of a Physical System Model", Journal of Dynamic Systems, Measurement, and Control, ASME, PP.185-186, 1979.
- [24] Karnopp, D. " Alternative Bond Graph Causal Patterns and Equation formulations for Dynamic Systems", Journal of Dynamic Systems, Measurement, and Control, ASME, PP.58-63, 1983.
- [25] Karnopp, D. " An Approach to Derivative Causality in Bond Graph Models of Mechanical Systems", Journal of Franklin Institute, PP.65-75, 1992.
- [26] Kokotovic, P. Khalil, H.K., and O'Reilly, J. "Singular perturbation Methods in Control, Analysis and Design", Academic Press, London, 1986.

- [27] Martens, H.R. "Simulation of Nonlinear Multiport Systems Using Bond Graphs", Journal of Dynamic Systems, Measurement, and Control, ASME, PP.49-54, 1973.
- [28] Joseph, B.J., Martens, H.R. "The Method of Relaxed Causality in the Bond Graph Analysis of Nonlinear Systems", Journal of Dynamic Systems, Measurement, and Control, ASME, PP.95-99, 1974.
- [29] Maschke, B.M., Yazman, M.A. "Graphical Tools to Analyze Nonlinear Controlled Systems : Bond Graphs and System Graphs", Proceedings of IMACS, Vol 1. PP.23-26, 1988.
- [30] Maschke, B.M., "Bond Graphs for the Structural Dynamic Decoupling Problem : A Case Study", Proceedings of IMACS, PP.1074-1078, Vol.3, 1991.
- [31] Miu, D.K., "Physical Interpretation of Transfer Function Zeros for Simple Control Systems with Mechanical Flexibilities", Journal of Dynamic Systems, Measurement and Control, Vol. 113, pp. 419-424, September 1991.
- [32] Nolan, P.J. "Symbolic Bond Graph Processing using LISP", Proceedings of IMACS World Congress on Computation and Applied Mathematics, 1991.
- [33] Nolan, P.J. "Symbolic and Algebraic Analysis of Bond Graphs", Journal of Franklin Institute, Vol.328, No. 5/6, pp.1027-1046, 1991.
- [34] Redfield, R.C.and Krishnan, S., "Dynamic System Synthesis with a Bond Graph Approach: Part I - Synthesis of One-Port Impedances" Journal of Dynamic Systems, Measurement, and Control Vol.115, Sep, 1993.

- [35] Redfield, R.C.and Krishnan, S., "Dynamic System Synthesis with a Bond Graph Approach: Part II - Conceptual Design of an Inertial Velocity Indicator" Journal of Dynamic Systems, Measurement, and Control Vol.115, Sep, 1993.
- [36] Rosenberg, R.C. " State-Space formulation for Bond Graph Models of Multiport Systems", Journal of Dynamic Systems, Measurement, and Control, ASME, PP.35-40, 1971.
- [37] Rosenberg, R.C., Beaman, J. " Clarifying Energy Storage field Structure in Dynamic Systems", Proceedings of ACC,pp.1451-1456, 1987.
- [38] Rudin, W. " Principles of Mathematical Analysis", McGraw-Hill, Inc., 1976.
- [39] Slotine, J-J.E., Li, WeiPing, "Applied Nonlinear Control", Prentice-Hall, 1991.
- [40] Strang, G. "Introduction to Applied mathematics", Wellesley-Cambridge Press,1986.
- [41] Suda, N., and Hatanaka, T., "Structural Properties of Systems Represented by Bond Graphs", , Complex and Distributed Systems: Analysis, Simulation and Control, Congress IMACS 1986.
- [42] SUDA, N. " Effect of Parasitic Storage Elements in Systems Represented by Bond Graphs", IMACS 1988 Proceedings of the 12th World Congress on Scientific Computation, Vol.1, pp.17-19, 1988.
- [43] Sueur, C. and Dauphin-Tanguy, G., "Structural Controllability/Observability of Linear Systems Represented by Bond Graphs" J. Franklin Institute, Vol.326, No.6, pp.869-883, 1989.
- [44] Sueur, C. and Dauphin-Tanguy, G., "Bond Graph Approach for Structural Analysis of MIMO Linear Systems," J. Franklin Institute, Vol.328, pp.55-70, 1991.

- [45] Sueur, C. and Dauphin-Tanguy, G., "Bond Graph Approach for Multi-time Scale System Analysis" J. Franklin Institute, Vol.328, No.5/6, pp.1005-1026, 1991.
- [46] Van Dijk, J., Breedveld, P.C. " Models Containing Zero-Order Causal Paths-I,II. Classification of Zero -order Causal Paths", Journal of franklin Institute, 1992.
- [47] Wu, S.-T. and Youcef-Toumi, K. "On Relative Degrees and Zero Dynamics from System Configuration", Journal of Dynamic Systems, Measurement and Control, ASME, 1995.
- [48] Wu, S.-T., "Input/Output Linearization of Uncertain Systems with Time Delay Control", PhD. Thesis, Department of Mechanical Engineering, M.I.T., 1993.
- [49] Yoshimura, H., Nakano, H., and Kawase, T. "Flexible Multibody Dynamics and Symbolic Generation Scheme", Proceedings of IMACS World Congress on Computation and Applied Mathematics, 1991.
- [50] Zeid, A. " An Algorithm for Eliminating Derivative Causality", Proceedings of ASME 1988 Winter Annual Meeting, DSC-Vol.8, pp. 15-22, 1988.
- [51] Zeid, A. and Rosenberg, R., "Estimating Eigenvalues for a Class of Dynamic Systems", J. Franklin Inst., pp.21-40, July 1985.
- [52] Zeid, A. and Rosenberg, R., "Eigenvalue Spectra and Bounds for Certain Classes of Dynamic Systems Having Tree Bond Graphs", IEEE Transactions on Circuits and Systems, Vol. CAS-33, No.12, December 1986.
- [53] Zeid, A. "Some Bond Graph Structural Properties: eigen spectra and stability", Proceedings of ASME Winter Annual Meeting, PP.75-82, 1988.