REALIZATION-PRESERVING SIMPLIFICATION AND REDUCTION OF DYNAMIC SYSTEM MODELS AT THE GRAPH LEVEL

by

Tulga Ersal

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Mechanical Engineering)
in The University of Michigan
2007

Doctoral Committee:

Professor Jeffrey L. Stein, Chair
Professor Gregory M. Hulbert
Professor Pierre T. Kabamba
Associate Professor Bogdan Epureanu
Assistant Research Scientist Hosam K. Fathy

To my parents

ACKNOWLEDGMENTS

I would like to give my deep gratitude to my committee chair, Prof. Jeffrey L. Stein, for his profound mentorship, intellectual guidance, endless patience, and support. I am also indebted to the rest of my committee for their insightful feedback. Specifically, I would like to thank Prof. Hulbert for his feedback on the coordinate frame reorientation part of this work; and Prof. Kabamba for his feedback on the model reduction part. I especially would like to thank Prof. Epureanu for graciously taking time to get me started on the Karhunen-Loève expansion and to discuss with me numerous times the subtleties of that part of this work that is based on the Karhunen-Loève expansion. I am also forever thankful to Dr. Hosam K. Fathy for first being my friend, and then my mentor. His attention to details, constructive criticism, and amazing ability to put things into perspective greatly improved the quality of this work. I cannot thank him enough for our immensely enjoyable and endless discussions, for being there whenever I needed him, and for believing in me even when I did not.

I would also like to thank Burit Kittirungsi for spending an enormous amount of time with me to develop the HMMWV model. Without his insight, resourcefulness, and willingness to help, I do not even want to think how long it would take me to finish the model. I am also thankful to Steve Riley for his interest in this work and his feedback, and for sharing his literature on multibody dynamics with me.

This research has been partially supported by the Department of Mechanical Engineering at the University of Michigan through fellowships, and also by the Engineering Research Center for Reconfigurable Manufacturing Systems of the National

Science Foundation under Award Number EEC 9529125. These supports are gratefully acknowledged.

Most of the chapters of this dissertation are heavily based on manuscripts that have been published or have been submitted for publication. Specifically, Chapter 2 is based on [1], Chapter 3 on [2, 3], Chapter 4 on [4, 5], and Chapter 5 on [6, 7]. Moreover, even though no separate chapter is devoted to it, the bond-graph based modular modeling environment utilized throughout this work is based on [8-10]. I gratefully acknowledge the contributions of my co-authors on those publications.

Finally, I would like to thank my parents, Dilek and Bora Ersal, for warming my heart and easing my mind despite the great distance between us as only they can do; and my wife, Ilkin, for sharing every aspect of this experience with me, for her patience, tolerance, and love, and for making me believe.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

LIST OF APPENDICES

Appendix

CHAPTER 1

INTRODUCTION

"*Everything should be made as simple as possible, but not simpler*."
– Albert Einstein

1.1. PROBLEM SPACE

Even though physical experiments are still indispensable in engineering, computers provide powerful virtual alternatives, when physical experiments are too time-consuming, expensive, or even impossible to conduct [11]. These virtual experiments, or simulations, prove invaluable for developing and testing new design concepts rapidly and inexpensively, and give vital insight into the system behavior. They rely on mathematical models that describe physical phenomena; hence, they are only as successful as the underlying mathematical model in faithfully and efficiently recreating physical systems in the virtual environment. Therefore, mathematical models are at the heart of simulation-based engineering and science.

To maximize their utility, mathematical models need to have two fundamental characteristics: fidelity and simplicity. For a model to be reliable, it is critical that it has enough fidelity, i.e., it describes the behavior of the system it represents accurately enough for the intended purpose. It is also important that the model is simple enough to comprehend, handle, and simulate efficiently. In other words, a model has to be only as complex as necessary to fulfill its purpose.

Achieving fidelity and simplicity simultaneously, however, is challenging, because these two characteristics typically translate into conflicting targets. As more

physical phenomena are included in the model, the model's fidelity increases, but its simplicity starts to suffer. If, on the other hand, too few phenomena are incorporated into the model to promote simplicity, the fidelity is impaired. Thus, it is important to find a balance between fidelity and simplicity.

This balance can only be achieved by capturing those phenomena, and only those phenomena that dominate the system's behavior in the scenario of interest. Thus, it is critical to know which phenomena are crucial to include in the model, which phenomena can be neglected, and which phenomena are irrelevant for the considered scenario. However, this valuable knowledge is difficult to have, as it requires much time and expertise. Therefore, balancing fidelity and simplicity is a great challenge.

1.2. SCOPE OF THE WORK

A model that balances fidelity and simplicity is deemed *proper* in literature [12]. Working with proper models is highly desired, yet obtaining them is challenging because of the aforementioned reasons. Therefore, many tools have been developed to aid the modeler in this task. These tools can be divided into two major categories based on their approach to proper modeling:

1. Deduction tools: These tools assume that there exists a baseline model that captures only some basic phenomena, is simple, but lacks fidelity. Then, they systematically look for and add next conceivably most important phenomena to the model until it becomes proper.

2. Reduction tools: These tools assume that there exists a baseline model that captures more phenomena than necessary, and has therefore excessive fidelity and lacks simplicity. Then, they systematically look for and eliminate conceivably negligible phenomena from the model until it becomes proper.

Between these two approaches reduction is the more common one in the literature. This means that initially preference is given to fidelity over simplicity to satisfy the fidelity requirement first. Then, reduction tools are utilized to decrease the complexity of the model with an acceptable compromise in fidelity to achieve simplicity. This philosophy is also accepted in this work.

This work distinguishes between two special cases of reduction:

1. Simplification: This refers to eliminating from the model only those phenomena whose elimination does not affect model fidelity. In other words, simplification removes only the irrelevant phenomena, and can be considered as a reduction with 100% fidelity retention (within a numerical tolerance). For example, the well-known ideas of pole-zero cancellation [13], Kalman's minimal realization [14], or explicit elimination of Lagrange multipliers [15, 16] can be referred to as simplification techniques.

2. Partitioning: This refers to breaking weak two-way couplings in a model into one-way connections to create driving and driven submodels. When this is done, the output of the driving submodel acts like an input to the driven submodel, but the driven submodel does not affect the driving submodel. Models can be partitioned, e.g., into slow and fast dynamics [17], high- and low-frequency oscillation modes [18], heavily- and lightly-damped dynamics [18], or driven and driving sections [19].

To illustrate the difference between partitioning, simplification and reduction in general, consider an automobile that is modeled with 6 DoF (degrees of freedom), but moves only on the pitch plane. Then, assuming that the translational dynamics affect the pitch dynamics, but not vice versa, decoupling the two would be deemed as partitioning. Recognizing that the vehicle motion is on a plane and eliminating yaw, roll, and lateral

dynamics to end up with a 3 DoF model would be simplification. Finally, if the pitch dynamics are considered unimportant for the system dynamics of interest and removed from the model completely, that is considered reduction in a more general sense. It is worth noting that these distinctions are not necessarily universally accepted, but are used in this work for ease of presentation.

This work deals in particular with simplification and reduction in general, but partitioning is out of scope.

As for the types of systems considered in this work, the framework can be described as follows: First of all, this work concentrates on energetic systems only, i.e., systems composed of components that store, dissipate and exchange energy. Since such systems are quite common in engineering, this focus does not present a severe limitation. The motivation behind this focus is to take advantage of the domain-independent and intuitive notions of energy and power in simplification and reduction considerations.

Furthermore, only deterministic systems will be considered, i.e., it will be assumed that a system's future state is completely determined by its current state and current and future inputs without any random effects.

Finally, only lumped-parameter representations of systems are considered. This is also not considered a severe limitation, as this is a typical way of representation in the area of system dynamics and control. Moreover, distributed systems can be brought into the scope of this work through their lumped-parameter approximations.

Therefore, in summary, the scope of this work is the simplification and reduction of lumped parameter models of nonlinear energetic deterministic systems.

1.3. PROBLEM STATEMENT

A literature survey on proper modeling techniques is given the next chapter. The conclusion of the survey is that each of the existing methods has one or more of the following limitations:

1. Applicability to a limited set of systems: Some proper modeling techniques are limited to particular classes of systems (e.g., linear systems, time-invariant systems, etc.).

2. Requiring a realization change: Many proper modeling techniques project the dynamics of the given system onto a new state space conducive to simplification/reduction. This is often attractive in terms of minimizing the approximation error, but one may conceivably model a system using particular state variables and wish for this realization to remain invariant during simplification/reduction, perhaps because of its intuitive appeal. Any simplification/reduction technique that meets this need is referred to as realization-preserving in this work.

3. Trajectory independence: Many proper modeling techniques seek models whose accuracy is acceptable over a broad range of state and input trajectories. Such input independence is often attractive, but one may conceivably seek a model that is proper only for a given trajectory or small family of trajectories. For instance, a vehicle safety engineer may seek a simple model that accurately captures vehicle dynamics only over the set of all maneuvers likely to induce rollover. In such situations, trajectory-independent proper modeling algorithms may furnish excessively complex models, and trajectory-dependent approaches may be preferable.

4. Being limited to equation level: Graph representations (e.g., linear graphs, bond graphs, etc.) often provide intuitively appealing depictions of system models, but most proper modeling algorithms operate at the equation level. This means that even though the equations derived from a graph representation could be simplified/reduced using the existing techniques, the simplification/reduction would not necessarily be reflected at the graph level, which may hinder the advantages of having a graph-level representation.

5. Not considering the structure of the model: Most proper modeling methods seek to simplify/reduce the *order* of a given model, i.e., its number of states. Therefore, possible simplifications/reductions in the *structure* of the model, i.e., how the components of the given system interact with each other, are typically not taken into account.

This work strives to develop model simplification and reduction algorithms that address the above limitations. Specifically, algorithms are sought that are realization-preserving, trajectory-dependent, applicable to graph level representations of nonlinear systems, and also aimed at both structure and order simplification/reduction.

1.4. HYPOTHESES

Because of the fact that the energy flow patterns in an energetic system determine the system's behavior, this work proposes to use energy as the basis of the simplification and reduction algorithms to be developed. Such energy-based methods would allow for a unified treatment of not only different energy domains (e.g., mechanical, electrical, hydraulic, etc.), but also the dynamic components and their interactions in a system, thereby enabling simultaneous order and structure simplification/reduction.

Specifically, the hypotheses of this work can be summarized as follows:

1. For simplification purposes, a weighted $\mathcal{L}_1$-norm of power known as activity in literature [20] can be used to detect and eliminate the structure that has no effect on the fidelity of the model. In particular, the junction elements (structural components in a model that implement Kirchhoff's generalized node and loop laws) whose connections have almost zero activity can be removed from the model without compromising fidelity.

2. For more general reduction purposes, taking into account not only the magnitude, but also the correlations between the energy flow patterns throughout the model would allow for a better assessment of the relative importance of each part of the model to the system behavior. In addition, such an evaluation could also be extended to the structure of the model to make simultaneous order and structure possible.

1.5. MODELING PHILOSOPHY AND LANGUAGE

A modular approach to modeling is adopted here to efficiently create the initial models. This means that the system models are obtained from generic component models through an assembly process. Due to reuse of submodels, the modular approach allows for efficient building, verification and handling of large-scale systems.

Although models can be obtained rapidly through the modular approach, they are hardly proper at the outset. This is because of the level of detail included in the generic component models to promote modularity. For example, the generic rigid body model used in this work is created with 6 DoF, but in most engineering applications components hardly retain all 6 DoF when assembled into a system. Therefore, the modular model is subsequently simplified or reduced to make it proper.

Therefore, in summary, the modeling philosophy of this work is to model the system in a modular way first, and then simplify/reduce it to obtain a proper model. It is

argued that with the aid of effective simplification/reduction tools this philosophy would lend itself very well to automated proper modeling of dynamic systems.

With the modeling philosophy being modular and the scope being energetic systems, it is found especially convenient to use in this work the bond graph language to represent models due to its modular, power-based nature [21]. Furthermore, this language provides visual, intuitive, and compact representations of models, which will be important for presentation purposes especially in Chapter 6, where the proper modeling of a large-scale system is discussed. It is worthwhile to emphasize in advance, however, that none of the methods developed in this work is fundamentally limited to bond graphs. Nevertheless, they are easily implemented and presented using bond graphs.

1.6. ORGANIZATION OF THIS DISSERTATION

The rest of this dissertation is organized as follows. First, Chapter 2 reviews the state of the art in proper modeling. Chapter 3 formally introduces the inactivity of a junction element, and presents the proposed inactive-junction-based simplification algorithm. Chapter 4 proposes a method to relax the realization-preserving property of the simplification algorithm to the extent that body-fixed coordinate frames in multibody systems are automatically reoriented to achieve better simplification results, while still preserving the intuitive appeal of the model. Chapter 5 introduces a new energy-based metric to evaluate the importance of the various dynamic and structural parts of a model. It then presents the proposed reduction algorithm based on this metric, which enables simultaneous order and structure reduction. Chapter 6 is a case study which presents the proper modeling of the Army's High Mobility Multipurpose Wheeled Vehicle, HMMWV. Chapter 7 concludes this dissertation with a summary, and lists of contributions and possible extensions of this work.

CHAPTER 2

A REVIEW OF THE PROPER MODELING LITERATURE

2.1. INTRODUCTION

This chapter briefly surveys the numerous proper modeling techniques presented in the literature. Some of these techniques begin with simple models and increment their complexity until they meet their respective accuracy requirements: a process known as *model deduction*. Most techniques, however, begin with excessively complex models and then *reduce* them until they become proper. The ultimate goal of both model deduction and reduction techniques is the same, regardless of how it is achieved: given a dynamic system model, balance its accuracy and complexity by massaging it to include only the most salient dynamics of the given system.

This implies that every proper modeling algorithm must have at its core a *metric* for quantifying the relative importance of modeling the different dynamics of a given system. Based on the metrics they use for proper modeling, this chapter classifies the proper modeling techniques presented in the literature into *frequency*-, *projection*-, *optimization*-, and *energy-based*. This classification is neither a universally adopted convention, nor is it strict. In fact, the chapter shows that a given proper modeling technique can often conceptually belong to more than one of these categories. However, the author finds this classification intuitively appealing and convenient for presentation, and hence adopts it herein.

Similar reviews exist in the literature [22-32], but this review presented in this chapter is unique in its use of proper modeling as a broad contextual framework within which different algorithms are compared and contrasted.

2.2. FREQUENCY-BASED TECHNIQUES

The fundamental metric used by *frequency-based* proper modeling techniques for assessing the importance of a given system's various dynamics is *characteristic speed*. In particular, given a dynamic system model, these techniques partition it into submodels with comparatively "fast" and "slow" dynamics whose relative importance depends on the given application.

Consider, for instance, the dynamics of a hydraulic car braking system. A full model of such a system may simultaneously capture the dynamics of the car's motion and the dynamics of hydraulic pressure wave propagation. The latter dynamics are typically orders of magnitude faster than the former. A model capturing both sets of dynamics is therefore likely to exhibit significant *numerical stiffness*, defined as a disparity between its different characteristic speeds. Such numerical stiffness may cause the model to be computationally intractable, thereby necessitating a more "proper" technique for modeling this braking system. Such a proper modeling technique may neglect fluid compressibility when the goal is to examine vehicle braking, and conversely neglect vehicle motion when the goal is to examine pressure wave propagation.

This chapter refers to all techniques that use characteristic speed as a metric for proper modeling as *frequency-based* techniques. The term "frequency-based", in this context, underscores the congruence between characteristic speeds and eigenvalues in the case of linear systems. Indeed, as the review below shows, frequency-based proper modeling techniques are most often used for linear systems, even though many of them can be generalized to nonlinear systems. This review focuses on eight established classes

of frequency-based proper modeling techniques from the literature, namely, *aggregation*, *singular perturbation*, *the model order deduction algorithm (MODA)*, *modal analysis*, *component mode synthesis (CMS)*, *polynomial approximation methods*, *oblique projection*, and *optimal Hankel norm approximation*. It briefly details the fundamental principles behind each technique or class of techniques, in addition to their conceptual similarities and differences.

2.2.1. Aggregation

One of the basic ideas in the model reduction literature is to ignore the small time constants in a system, and keep the large ones, which dominate the response. Thus, the earlier model reduction methods were based on retaining the dominant eigenvalues of the system in the reduced model [33-40]. While developing his optimal projection method Mitra showed that Davison's method [33] is a special case of optimal projection [41, 42]. Aoki later developed the more general method of aggregation [43], and it has been shown that Mitra's optimal projection method is a special case of aggregation [44-46].

The basic idea behind the aggregation method can be summarized as follows. Consider the approximation of the *n*-dimensional original system

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx + Du \end{aligned} \tag{2.1}$$

with the *r*-dimensional reduced model

$$\begin{aligned} \dot{x}_r &= A_r x_r + B_r u \\ y &= C_r x_r + Du \end{aligned} \tag{2.2}$$

Suppose the reduced state vector x_r is related to the original state vector *x* through

$$x_r = Kx \tag{2.3}$$

where *K* is the $r \times n$ aggregation matrix. It follows that

$$\begin{aligned} A_r K &= KA \\ B_r &= KB \\ C_r K &\approx C \end{aligned} \tag{2.4}$$

A least-squares solution can be obtained by using the pseudoinverse as

$$\begin{aligned} A_r &= KAK^{\dagger} \\ B_r &= KB \\ C_r &= CK^{\dagger} \end{aligned} \tag{2.5}$$

It has been shown that a nontrivial aggregation law exists if and only if the A_r retains r of the eigenvalues of A [46]. Furthermore, K can be obtained by

$$K = T\begin{bmatrix} I_r & 0 \end{bmatrix} V^{-1} \tag{2.6}$$

where T is any nonsingular matrix, I_r is the $r \times r$ identity matrix and V is the modal matrix of A.

This basic idea of aggregation has been extended by many researchers. For example, Aoki proposed two ways of relaxing the perfect-aggregation condition [47]. Hickin proposed a method called nonminimal partial realization that combines the ideas of aggregation and moment matching [48]. Siret et al. developed a method to chose the arbitrary matrix T in Eq. (2.6) in an optimal way to maximize a performance criterion [45]. It must be noted, however, that even though some of the eigenvalues of A are retained, the aggregation method is not realization-preserving, because the reduced model uses a different set of state variables than the original one; specifically, a combination of the original state variables. Hence, the intuitive appeal of the original model may not be preserved in the reduced model.

2.2.2. Singular Perturbation Method

As the difference between the large and small time constants in a system increases, or, in other words, as the underlying characteristic speeds become significantly disparate, the system is said to possess multiple time scales. In this case, if the interest is in the slow time scale, the problem may become numerically stiff due to the fast time scale. *Singular perturbation* is a reduction technique particularly suited to this type of models.

Unlike aggregation, singular perturbation is realization-preserving in the sense that it does not necessarily require a coordinate transformation as part of model reduction. This is quite attractive, because it implies that the physical meaning associated with each state in the original model can be preserved in the reduced model.

In its simplest rendition, singular perturbation implicitly assumes *a priori* knowledge of which state variables of a given model correspond to the fast dynamics and which correspond to the slow. Neglecting the influence of the "fast" dynamics on the "slow" states *partitions* the original stiff model into two submodels. The first *driving* submodel captures the slow dynamics and *residualizes* the fast states, while the second *driven* submodel captures the fast dynamics and treats the slow states as input variables. This furnishes a *decoupled* system model that not only mitigates the original model's numerical stiffness but also approaches the original model in accuracy as this stiffness grows.

Singular perturbation is *realization-preserving* in the sense that it does not necessarily require a coordinate transformation as part of model reduction. This is quite attractive, because it implies that the physical meaning associated with each state in the original model can be preserved in the reduced model.

The origins of the singular perturbation method go back to Prandtl's work on boundary layers in fluid dynamics [49]. Later contributions by Tikhonov [50], Levinson [51], Vasileva [52], Wasow [53] and Kokotovic [54-57] established singular perturbation as a model reduction tool. In its simplest rendition, the singular perturbation method assumes that the dynamics of a system are expressed in state space form, where some derivatives have a small positive number ε as a coefficient, i.e.,

$$\dot{x}_1 = f_1(x_1, x_2, u), \quad x_1 \in \mathbb{R}^n \tag{2.7}$$

$$\varepsilon \dot{x}_2 = f_2(x_1, x_2, u), \quad x_2 \in \mathbb{R}^m \tag{2.8}$$

The coefficient ε represents the *disparity* between the characteristic speeds of the fast and slow dynamics. As this coefficient approaches zero, Eq. (2.8) becomes

$$0 = f_2(\bar{x}_1, \bar{x}_2, \bar{u}) \tag{2.9}$$

where bars are used to distinguish between this limiting case and the case where ε truly equals zero. Now assume that Eq. (2.9) can be solved to obtain a distinct real expression for $\bar{x}_2$ in terms of $\bar{x}_1$, i.e.,

$$\bar{x}_2 = \phi(\bar{x}_1, \bar{u}) \tag{2.10}$$

Substituting this solution into Eq. (2.7) effectively furnishes a slow submodel that *residualizes* the fast states, i.e.,

$$\dot{\bar{x}}_1 = f_1(\bar{x}_1, \phi(\bar{x}_1, \bar{u}), \bar{u}) = \bar{f}(\bar{x}_1, \bar{u}) \tag{2.11}$$

The reduced model for the fast dynamics can be obtained by introducing a fast time scale τ and fast variables $\tilde{x}_1(\tau)$ and $\tilde{x}_2(\tau)$ defined as follows:

$$\tau = \frac{t}{\varepsilon}, \quad x_j(t) = \bar{x}_j(t) + \tilde{x}_j(\tau), \quad j = 1, 2 \tag{2.12}$$

Combining Eq. (2.7), (2.8), and (2.12), and letting $\varepsilon \to 0$, the fast-dynamics model is obtained as

$$\frac{d\tilde{x}_2}{d\tau} = f_2(x_1(t), \bar{x}_2(t) + \tilde{x}_2(\tau), u) \tag{2.13}$$

This model uses the slow states as inputs, and is hence *driven* by them.

Equations (2.7)-(2.13) highlight the simplicity with which the singular perturbation method can be applied to a given system. In addition to this simplicity and the method's intuitive appeal, the singular perturbation method furnishes reduced models with attractive mathematical properties in some special cases. In particular, let the original and reduced models be G and G_r, respectively. Furthermore, assume that the full model G is expressed in the time domain using a *balanced realization* (see Section 2.3.2), then reduced to G_r using the singular perturbation method. Then, the singular perturbation method is equivalent to *balanced residualization*, a projection-based proper modeling technique. Furthermore, the maximum error introduced by singular perturbation, quantified in terms of the $\mathcal{H}_\infty$ norm of the difference $G - G_r$, satisfies:

$$\left\| G - G_r \right\|_\infty \leq 2(\sigma_{n+1} + \ldots + \sigma_{n+m}) \tag{2.14}$$

where σ_i, $i = n+1,...,n+m$ are the Hankel singular values of G corresponding to the fast dynamics [58]. In other words, the $\mathcal{H}_\infty$ norm of the modeling error introduced by singular perturbation cannot exceed twice the sum of the Hankel singular values corresponding to the fast states. Furthermore, this modeling error decreases with the parameter ε, and becomes zero in the limit as ε approaches zero.

2.2.3. Model Order Deduction Algorithm

Like singular perturbation, the *model order deduction algorithm (MODA)* is a realization-preserving technique that deems a model "proper" if it captures only the most relevant characteristic speeds of a given system for a given application. Unlike singular perturbation, however, MODA is a *deduction* algorithm that starts with simple models and increments their complexity until they become proper. Furthermore, MODA does not assume *a priori* knowledge of which states in a system are "fast" and which are "slow". Instead, it explicitly searches for this knowledge as part of its pursuit of proper models.

In its simplest rendition [59], MODA deems a linear system model proper for a given application if the model's *rank* is minimal and its *spectral radius* exceeds a *frequency range of interest (FROI)* desired for the application. The spectral radius of a linear system is defined as the radius of a closed ball containing all its poles, or equivalently, as the Euclidian norm of its largest poles. Furthermore, the rank of a model, in this context, is the number of components in the model not included in the initial baseline model from which the deduction process proceeds. For instance, a finite-element model of a shaft that uses 30 finite elements has a rank of 23 compared to a baseline finite element model of the same shaft that uses only 7 finite elements.

MODA begins with a baseline model and proceeds to increment its rank in a manner that produces the smallest increase in its spectral radius, repeating this process until the spectral radius exceeds the desired FROI [59]. Using this approach, MODA

furnishes not only a proper model, but also an understanding of which subsystem dynamics need to be captured accurately to furnish a proper system model. For instance, given a system containing more than one flexible shaft, MODA can determine the number of finite elements needed to model each shaft so that the overall system model is proper. This makes MODA particularly attractive for the automated lumped-parameter modeling of continuous dynamic systems [59].

The literature describes several extensions that enhance the capabilities of MODA. In particular, Ferris *et al.* extend MODA to not only satisfy a given spectral-radius requirement, but also capture system eigenvalues within that spectral radius with a desired level of accuracy [60]. Furthermore, Walker *et al.* modify this algorithm to furnish models that accurately capture the eigenvalues of only the observable and controllable modes of a given system within the desired FROI [61]. Wilson and Taylor modify MODA to seek an accurate representation of a system's frequency response within the desired FROI as opposed to just its eigenvalues [62]. Finally, Taylor and Wilson extend MODA to enable the proper modeling of nonlinear systems over a desired range of input excitation frequencies [63].

MODA is not the only algorithm that adopts the deduction approach to proper modeling. Pirvu *et al.*, for example, propose a bond-graph-model adaptation algorithm that searches for all possible extensions of a given baseline bond-graph model that would result in a desired higher-order transfer function [64]. The baseline model can be extended by adding new interconnections, i.e., 1- and 0-junctions in bond graph terms, or energetic components, i.e., generalized inductors, capacitances or resistors. The transfer-function-matching objective, however, limits this method to linear systems.

Another example of the deduction approach is the bond-graph synthesis using genetic algorithms [65, 66]. Similar to Pirvu's method, this method lets a bond graph evolve from a baseline model. However, the freedom in choosing the fitness function

gives this method more flexibility, allowing it to be used not only as a proper modeling tool, but also a conceptual system synthesis tool.

2.2.4. Modal Analysis

In its simplest rendition, *modal analysis* focuses on linear, time-invariant, vector-second-order dynamic systems satisfying the principle of separation of variables (e.g., through proportional damping). Such systems may be finite- or infinite-dimensional. In the latter case, one often approximates the given system's continuous dynamics using a finite-dimensional, lumped-parameter model obtained through a discretization technique (such as finite differences or finite elements). The resulting finite-dimensional model of this vector-second-order system, subject to the assumption of negligible damping, can be expressed as [67, 68]

$$M\ddot{x} + Kx = 0 \tag{2.15}$$

where M and K are the effective structural inertia and stiffness matrices, respectively. The modes of such a system can be found by solving the generalized eigenvalue problem

$$Kv = \omega^2 Mv \tag{2.16}$$

where the natural frequencies are given by the various solutions for ω and the modes shapes are given by the corresponding solutions for v. These mode shapes collectively form a basis spanning the complete state space corresponding to Eq. (2.15). Therefore, the dynamics represented by Eq. (2.15) can be projected onto the eigenspace given by these mode shapes without loss of information. Such a projection can also be performed on the standard state-space representation of the full model (as opposed to the vector-second-order representation), leading to a new state-space model with a diagonal A matrix (with complex entries), as shown below:

$$\dot{\xi} = A\xi + Bu,\ \ y = C\xi$$

$$A = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix},\ B = \begin{bmatrix} b_1^T \\ b_2^T \\ \dots \\ b_n^T \end{bmatrix},\ C = \begin{bmatrix} c_1 & c_2 & \dots & c_n \end{bmatrix} \tag{2.17}$$

Given this new *modal* representation, modal analysis builds on the congruence between the eigenvalues corresponding to a given mode and the characteristic speed of the mode to achieve model reduction. In particular, by eliminating the modes corresponding to the faster eigenvalues from Eq. (2.17), one can balance the fidelity and complexity of a given model, thereby rendering it proper [67, 68]. Modal analysis is therefore a frequency-based model reduction technique that does not assume *a priori* knowledge of which dynamics of a given system are "fast" and which are "slow". Like singular perturbation, it has the very attractive property of a guaranteed error bound. In particular, the $\mathcal{H}_\infty$ norm of the difference between the original model, G, and reduced model, G_r, is bounded by

$$\left\| G - G_r \right\|_\infty \le \sum_{i=k+1}^{n} \bar{\sigma}\left(c_i b_i^T \right) \Big/ \mathrm{Re}\left[\lambda_i \right]. \tag{2.18}$$

where λ_i is the i^{th} eigenvalue, and $\bar{\sigma}$ is the largest singular value of the residues $c_i b_i^T$ [69]. Unlike singular perturbation and MODA, however, modal analysis is not realization-preserving. It expresses the reduced model in terms of modal – rather than physical – coordinates. Consequently, physical insights associated with the original coordinate choice may be lost. Modal analysis shares this property with all *projection-based* proper modeling techniques, and is hence both a frequency-based and projection-based model reduction technique.

The simple rendition of modal analysis presented above only applies to linear finite-dimensional systems. There are several important extensions of this technique, however, that make it applicable to a broader range of problems. First, modal analysis can be applied directly to the partial differential equations governing the dynamics of an

infinite-dimensional system: a process that can furnish proper lumped-parameter models of such systems directly. Furthermore, the literature presents many extensions of modal analysis to both linear and nonlinear deterministic and stochastic systems that do not satisfy the assumptions of the above discussion [70-72]. Finally, the literature describes a special extension of modal analysis to modular systems known as *component mode synthesis*. This extension is discussed in further detail below.

2.2.5. Component Mode Synthesis

Component mode synthesis is an extension of modal analysis that is particularly applicable to large, modular systems. It proceeds in two simple steps. First, it uses modal analysis to obtain a proper model of each module in the system separately. Then it assembles these proper module models into a system-level proper model. This two-step approach can be significantly less expensive from a computational standpoint than the direct application of modal analysis to the entire system model, because solving many small eigenvalue problems can be significantly more tractable than solving one large eigenvalue problem. Because of its computational attractiveness, component mode synthesis is widely used in the literature [73-78], particularly in the context of applications involving large modular systems, such as automotive vibration applications [79-81].

2.2.6. Polynomial Approximation Methods

All five proper modeling techniques presented hitherto deem a model proper if it captures the dynamics of a system at either the “fast” or “slow” end of the frequency spectrum accurately and with minimal complexity. It is not uncommon, however, for one to pursue an accurate model of a system over one or more intermediate frequency bands. When modeling automobile noise, vibrations, and harshness (NVH), for instance, one is

usually interested in vibration frequencies small enough to be perceptible but large enough to cause potential passenger discomfort or drivability issues.

Padé approximation is a frequency-based model reduction technique particularly suited to this class of problems. Given a complex model, it finds a lower-order approximation of the model by first constructing Laurent series expansions of the frequency responses of both models at one or more interpolation points. It then matches a small number of coefficients of these expansions to parameterize the reduced model.

In particular, let $G(s)$ represent the transfer function of the original – or "full" – model. Then its Laurent series expansion around some $s_0 \in \mathbb{C}$ is given by

$$G(s) = \sum_{k=0}^{\infty} a_k (s - s_0)^k \tag{2.19}$$

The goal is to find a lower order model with the transfer function

$$G_r(s) = \sum_{k=0}^{\infty} \hat{a}_k (s - s_0)^k \tag{2.20}$$

such that for a desired number $n \in \mathbb{N}_0$, the equalities $a_k = \hat{a}_k$, $k = 0,1,2,\ldots n,$ are satisfied. The coefficients $a_k, \hat{a}_k$, $k = 0,1,2,\ldots,$ are referred to as moments, and therefore this technique is also known as *moment matching*. When $s_0 = \infty$, the moments become the Markov parameters of the system, in which case the approximation problem can be solved using the Arnoldi procedure [82, 83] or the Lanczos procedure [84, 85]. When s_0 is arbitrary, the rational Krylov method [86, 87] can be used. It is also possible to use multiple interpolation points [85, 87].

Padé approximation is attractive when one seeks a good local approximation of a model around certain interpolation points in the frequency domain at a low computational cost. However, the stability of Padé approximants is, in general, not guaranteed, even if the models being approximated are stable. The literature describes some techniques that address this problem by extending Padé approximation to seek only stable reduced models [88]. Two other important limitations of Padé approximation remain even with these methods. First, there are no global error bounds for Padé approximants. Secondly,

Padé approximation, by virtue of its dependence on the Laurent series expansion, is not a realization-preserving technique.

The starting point for Padé approximation is a Laurent series expansion of the frequency response of a given "full" model. If the full model is expressed as a rational polynomial transfer function, one may choose to obtain a proper model by truncating the polynomials in this transfer function directly, rather than expanding it into a Laurent series then performing moment matching. *Continued fraction expansion* is a polynomial approximation technique particularly suited to this scenario [89-93]. In particular, it builds on the fact that a transfer function given by

$$G(s) = \frac{a_{21} + a_{22}s + \ldots + a_{2,n}s^{n-1}}{a_{11} + a_{12}s + \ldots + a_{1,n-1}s^{n}} \tag{2.21}$$

can be written in the following continuous fraction expansion form

$$G(s) = \cfrac{1}{h_1 + \cfrac{1}{\cfrac{h_2}{s} + \cfrac{1}{h_3 + \cfrac{1}{\cfrac{h_4}{s} + \ddots}}}} \tag{2.22}$$

with

$$h_i = \frac{a_{i,1}}{a_{i+1,1}}, \quad i = 1, \ldots, 2n \tag{2.23}$$

where the coefficients a_{i1} are the first elements of the rows of the table

$$\begin{matrix} a_{11} & a_{12} & a_{13} & \ldots \\ a_{21} & a_{22} & a_{23} & \ldots \\ a_{31} & a_{32} & \ldots & \\ a_{41} & \ldots & & \\ \ldots & & & \end{matrix}, \quad a_{jk} = a_{j-2,k+1} - \frac{a_{j-2,1}a_{j-1,k+1}}{a_{j-1,1}}, \quad j = 3, \ldots n+1; \quad k = 1, 2, \ldots \tag{2.24}$$

This particular expansion, known as the second Cauer form [93], is just one of the possible forms of continued fraction expansion. Given this expansion, a reduced transfer function of order *r* can be obtained by retaining the first 2*r* coefficients *h* and truncating the rest. This preserves the steady state component of the original transfer function [27].

Other forms that can be used for continued fraction expansion include the first and third Cauer forms and the Stieltjes form [27, 93].

The main drawback of the continued fraction expansion method in general is that, like Padé approximation, unstable reduced models can result from stable original models. The literature addresses this problem by proposing other polynomial approximation methods guaranteed to preserve model stability. One such method is *Routh approximation* [94], which is based on the fact that a transfer function given by

$$G(s) = \frac{b_{11} + b_{12}s + \ldots b_{1n}s^{n-1}}{a_{11} + a_{12}s + \ldots + a_{1,n+1}s^{n}} \tag{2.25}$$

can be put into a canonical form, known as the alpha-beta expansion, given by

$$G(s) = \beta_1 f_1(s) + \beta_2 f_1(s) f_2(s) + \ldots + \beta_n f_1(s) f_2(s) \ldots f_n(s), \tag{2.26}$$

where

$$\begin{aligned} f_1 &= \frac{1}{1+\alpha_1 s} \\ f_i(s) &= \cfrac{1}{\alpha_i s + \cfrac{1}{\alpha_{i+1} s + \cfrac{1}{\vdots \\ \alpha_{n-1} s + \cfrac{1}{\alpha_n s}}}}, \quad i = 2, \ldots n \end{aligned} \tag{2.27}$$

and the coefficients α_i and β_i are given by

$$\begin{aligned} \alpha_i &= \frac{a_{i,1}}{a_{i,2}}, \; i = 1, \ldots n \\ a_{i,j} &= \begin{cases} a_{i-1,j+1} & j \; odd \\ a_{i-1,j+1} - \alpha_{i-1} a_{i-1,j+2} & j \; even \end{cases}, \; i = 2, \ldots n \\ \beta_i &= \frac{b_{i,1}}{a_{i,2}}, \; i = 1, \ldots n \\ b_{i,j} &= \begin{cases} b_{i-1,j+1} & j \; odd \\ b_{i-1,j+1} - \beta_{i-1} a_{i-1,j+2} & j \; even \end{cases}, \; i = 2, \ldots n \end{aligned} \tag{2.28}$$

A reduced model of order *r* can then be obtained by

$$G_r(s) = \frac{1}{s} \hat{G}_r \left(\frac{1}{s} \right) \tag{2.29}$$

with

$$\hat{G}_r(s) = \beta_1 p_1(s) + \beta_r p_1(s) p_2(s) + \ldots + \beta_r p_1(s) p_2(s) \ldots p_r(s)$$
$$p_1(s) = f_1(s) \tag{2.30}$$
$$p_2(s) = \cfrac{1}{\alpha_i s + \cfrac{1}{\vdots \quad \cfrac{1}{\alpha_{r-1} s + \cfrac{1}{\alpha_r s}}}}, \quad i = 2, \ldots r$$

In addition to preserving stability, the Routh approximant also guarantees that the first *r* coefficients of the Taylor series expansions about $s = 0$ of the original and reduced models match. Furthermore, the impulse-response energies of Routh approximants converge monotonically to those of the original models, and the poles and zeros of the approximants approach the ones of the original model as *r* increases [94].

The literature describes other polynomial approximation methods that preserve stability, such as reduction based on *stability equations* [95]. Furthermore, the literature describes mixed methods that use different methods for approximating the numerator and denominator. These methods aim to resolve the instability problem of the Padé and continued fraction expansion methods, while matching some quantities of the original model. Typically, dominant pole retention or some other stability-preserving polynomial approximation method is used to calculate the denominator of the reduced model, while Padé or continued fraction expansion is used to determine the numerator. Some combinations include dominant pole retention and Padé approximation [33, 35, 36, 38], Routh stability criterion and Padé approximation [96], Routh array and Padé approximation [97, 98], stability equations and Padé approximation [99], and stability equations and continued fraction expansion [100]. Nevertheless, two drawbacks of the polynomial approximation methods in general still remains, namely, that all such methods are limited to linear systems, and they are not realization-preserving.

2.2.7. Oblique Projection

Even though this method is, as its name suggests, a projection-based method, due to its close relationship with the polynomial approximation methods it will be reviewed here. The relationship is in the sense that this method, using the oblique projection approach, gives a unified tool to simultaneously match high and low frequency moments of the transfer function, and high and low power moments of the power spectral density [101].

This method frames the model reduction problem as a projection of the original model

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx + Du \end{aligned} \tag{2.31}$$

into the reduced model

$$\begin{aligned} \dot{x}_r &= A_r x_r + B_r u \\ y &= C_r x_r + Du \end{aligned} \tag{2.32}$$

by $A_r = LAT$, $B_r = LB$, $C_r = CT$, and $LT = I$. Then if L and T are chosen such that

$$L = \mathcal{O}_{p,q-1}(C);\ T = XL^T(LXL^T)^{-1} \tag{2.33}$$

where

$$\mathcal{O}_{p,q}(C) \triangleq \begin{bmatrix} CA^{-p} \\ CA^{-p+1} \\ \vdots \\ CA^{q} \end{bmatrix} \tag{2.34}$$

and X is the controllability Grammian satisfying

$$AX + XA^T + BB^T = 0 \tag{2.35}$$

then the reduced order model will be asymptotically stable if and only if it is controllable, and it will match p low frequency moments

$$M_i(0) = CA^{-i}B,\ i = 1, \ldots, p \tag{2.36}$$

q high frequency moments

$$M_i(\infty) = CA^{i}B,\ i = 0, \ldots, q-1 \tag{2.37}$$

p low frequency power moments

$$R_{ii}(0) = CA^{-i}X(A^T)^{-i}C^T,\ i = 1,\ldots,p \tag{2.38}$$

and, q high frequency power moments

$$R_{ii}(\infty) = CA^{i}X(A^T)^{i}C^T,\ i = 0,\ldots,q-1 \tag{2.39}$$

This basic idea has been extended to controller reduction at selected frequency regions, and also to matching the impulse response at selected time regions [101]. Due to its projection-based approach, this method is not realization-preserving.

2.2.8. Optimal Hankel Norm Approximation

The methods discussed so far deal with *local* approximations of a given system's frequency response. On the one hand, aggregation, singular perturbation, MODA, modal analysis, and component mode synthesis typically aim to approximate the low-frequency behavior of a given system. On the other hand, polynomial approximation methods typically aim to approximate the frequency response of a given system around some frequencies of interest.

Further extending these ideas, one may also seek a good approximation to a system's entire frequency response. Such an approximation may minimize, say, the $\mathcal{H}_\infty$ norm of the error $G - G_r$ between the full and proper models, but the resulting $\mathcal{H}_\infty$ model reduction problem does not have a known analytic solution. If, instead, one uses the Hankel norm of the error $G - G_r$ to quantify the "properness" of the reduced model, then an analytical solution for the optimal proper model does exist, and the resulting proper modeling technique is known as the *optimal Hankel norm approximation* [102-105].

For a given, stable, linear, and time-invariant system, G, Hankel norm approximation seeks an optimal reduced model, G_r, whose order, k, is specified *a priori* by the modeler. The resulting optimal proper model minimizes the Hankel norm of the error $G - G_r$ over the set of all linear and time-invariant models of the desired order. This

highlights the implicit tradeoff between fidelity (measured by the Hankel norm of $G-G_r$) and complexity (measured by the order of G_r) that makes Hankel norm approximation a proper modeling method. Assuming that the state-space description of *G* is given by $\{A,B,C,D\}$, one possible way of finding G_r of order *k* is as follows [106]:

1. Calculate *P* and *Q*, the controllability and observability Grammians of the system *G*, respectively.

2. Calculate $E=QP-\sigma_{k+1}^2 I$, where $\sigma_{k+1}=\sqrt{\lambda_{k+1}(PQ)}$ is the k+1st Hankel singular value of *G*.

3. Find the singular value decomposition of *E*,
$$E=\begin{bmatrix}U_1 & U_2\end{bmatrix}\begin{bmatrix}\Sigma & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}V_1^T\\ V_2^T\end{bmatrix}$$

4. Apply the transformation
$$\begin{bmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{bmatrix}=\begin{bmatrix}U_1^T\\ U_2^T\end{bmatrix}\left(\sigma_{k+1}^2 A^T+QAP\right)\begin{bmatrix}V_1 & V_2\end{bmatrix}$$
$$\begin{bmatrix}B_1\\ B_2\end{bmatrix}=\begin{bmatrix}U_1^T\\ U_2^T\end{bmatrix}\begin{bmatrix}QB & -C^T\end{bmatrix}$$
$$\begin{bmatrix}C_1 & C_2\end{bmatrix}=\begin{bmatrix}CP\\ -\sigma_{k+1}B^T\end{bmatrix}\begin{bmatrix}V_1 & V_2\end{bmatrix}$$

5. Form the equivalent model
$$\tilde{G}=\begin{bmatrix}\tilde{A} & \tilde{B}\\ \tilde{C} & \tilde{D}\end{bmatrix}$$
$$=\begin{bmatrix}\Sigma^{-1}\left(A_{11}-A_{12}A_{22}^{\dagger}A_{21}\right) & \Sigma^{-1}\left(B_1-A_{12}A_{22}^{\dagger}B_2\right)\\ C_1-C_2A_{22}^{\dagger}A_{21} & D\end{bmatrix}$$

6. The equivalent model can be decomposed additively into a stable part G_- with *k* stable poles and an anti-stable part G_+ with all poles unstable, i.e.,

$\tilde{G} = G_- + G_+$. Then, G_- is the *k*-th order optimal Hankel norm approximation of the system *G*, i.e., $G_r = G_-$.

The Hankel norm of the approximation error of any *k*-th order system $\hat{G}_r$ is lower-bounded by $\left\| G - \hat{G}_r \right\|_H \geq \sigma_{k+1}(G)$, and the equality in the error bound is satisfied only by the optimal Hankel norm approximation G_r.

This minimization of error in terms of the Hankel norm comes at the expense of a change in realization due the transformations applied during the calculation of the reduced model. Therefore, the optimal Hankel norm approximation is not a realization-preserving method.

It is worth noting that even though the Hankel norm approximation does not optimize $\mathcal{H}_\infty$ norm of the error, there still exists an $\mathcal{H}_\infty$ error bound, as established first by Glover [105]

$$\left\| G(j\omega) - G_r(j\omega) \right\|_\infty \leq 2 \sum_{i=k+1}^{n} \sigma_i \tag{2.40}$$

It is important to note that the *D* matrix does not affect the Hankel optimality of the approximation, but it does affect the $\mathcal{H}_\infty$ norm of the error. It is possible to choose $\tilde{D}$ in such a way that upper-bound on the $\mathcal{H}_\infty$ norm of the error is cut in half, i.e.,

$$\left\| G(j\omega) - G_r(j\omega) - \tilde{D} \right\|_\infty \leq \sum_{i=k+1}^{n} \sigma_i \tag{2.41}$$

Please see [105] for the calculation of such a $\tilde{D}$.

The above results for continuous systems have also been extended to discrete-time systems [107-110], multivariable systems [104, 111], and nonminimal systems [106]. An efficient computation method for large scale systems is given in [112].

2.3. PROJECTION-BASED TECHNIQUES

The frequency-based proper modeling techniques discussed hitherto assume, in general, that the salient dynamics of a given system occur over a fairly limited range in the frequency domain. *Projection-based* techniques make a conceptually analogous

assumption in the state domain. Specifically, they all assume that the salient dynamics of a given system are limited to a portion of the system's entire state space. They search for this portion – or *subspace* – by searching for the basis vectors spanning it, and they differ in the ways they choose the basis vectors. This section presents three projection-based model reduction techniques, namely, *Karhunen-Loève expansion, balanced truncation,* and *component cost analysis*.

2.3.1. Karhunen-Loève Expansion

The *Karhunen-Loève expansion* [113, 114], also known as principal component analysis [115], the method of empirical orthogonal functions [116], proper orthogonal decomposition [117], singular value decomposition [118], empirical eigenfunction decomposition [119-121], or the method of quasi-harmonic modes [122], is a correlation analysis tool that is a key foundation for most projection-based proper modeling techniques. It appears to have been suggested independently by several scientists [123], e.g., Kosambi [124], Loève [125], Karhunen [114], Pougachev [126], and Obukhov [127]. It can be implemented in a numerically efficient manner using the *method of snapshots* [119-121], and has become widely popular in many fields including fluid dynamics [128], random variables [129], image processing [130], signal analysis [131], data compression [132], oceanography [133], and process identification and control [134].

Given observation data from either a physical system or its model, the Karhunen-Loève expansion finds a subspace that captures the dominant dynamics of this system. Specifically, it finds the orthogonal basis that optimally captures the energy of the observation signals in the least-squares sense. Selecting those basis vectors that capture the most observation signal energy furnishes a subspace that captures the dominant system dynamics. Projecting the system's model onto this subspace using the *Galerkin*

projection method then furnishes a reduced model that captures the original system's dominant dynamics. This process leads to a powerful model reduction technique.

For time-invariant finite-dimensional systems, the Karhunen-Loève expansion method can be applied as follows. Consider a system represented by a state space equation of the form

$$\dot{x} = f(x,u), \quad x \in \mathbb{R}^n \tag{2.42}$$

Assume that $m \geq n$ observations are made for each state and arranged in matrix form such that

$$A = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}_{m \times n} \tag{2.43}$$

Obtain the singular value decomposition of the matrix A, i.e.,

$$A = U \Sigma V^T \tag{2.44}$$

where $\Sigma = diag(\sigma_1, \sigma_2, \ldots, \sigma_n)_{m \times n}$ with $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0$. The columns of the orthogonal $n \times n$ matrix V form a basis of the state space, and the squares of the singular values provide a measure of how much signal energy is captured by each of these basis vectors. Assume that the last $n-k$ singular values are small, where $k < n$. Then, a reduced order model can be obtained by taking the first k columns of the V matrix, and projecting the state space onto the subspace spanned by those k vectors, i.e.,

$$\dot{x}_r = (V_k)^T f(V_k x_r, u), \quad \bar{x} = V_k x_r \tag{2.45}$$

where $\bar{x}$ is the approximation to the original state vector x. The motivation for using the first k columns as a basis for the reduced model is the fact that the rank k approximation $A_k = U_k \Sigma_k (V_k)^T$ to the original observation matrix A is optimal in a least squares sense. Here U_k and V_k denote the first k columns of the matrices U and V, respectively, and Σ_k denotes the leading $k \times k$ principal minor of the matrix Σ. This optimality is guaranteed for any value of k. Furthermore, an error bound exists for the approximation error $A - A_r$, which is given by

$$\left\| A - A_r \right\|_F = \sum_{i=k+1}^{n} \sigma_i^2 \tag{2.46}$$

where $\|\bullet\|_F$ denotes the Frobenius norm [123, 135]. Note, however, that the optimality and the error bound are valid only for the approximation to the observation matrix, and not for the reduced order model, i.e., no bound exists for $\|x-\bar{x}\|$ in general. In fact, unstable reduced models may result from stable original models. Nevertheless, this technique often yields good results and is widely used for model reduction due to its applicability to nonlinear systems as well. Furthermore, under certain conditions it may be possible to characterize error bounds, consider the effects of small perturbations on the observation matrix, provide error estimates and ranges of validity [136-140].

In case the state variable is a function of position and time, $z(x,t)$, which is common in fluid mechanics or in structural vibrations, the same technique can be used to obtain empirical modes, such that the state variable can be approximated as

$$z(x,t) \approx \sum_{i=1}^{M} a_i(x) b_i(t) \tag{2.47}$$

In this case the observation matrix can be arranged as:

$$A = \begin{bmatrix} z(x_1,t_1) & z(x_2,t_1) & \cdots & z(x_n,t_1) \\ \vdots & \vdots & \vdots & \vdots \\ z(x_1,t_m) & z(x_2,t_m) & \cdots & z(x_n,t_m) \end{bmatrix}_{m\times n} \tag{2.48}$$

Then, the columns of the U matrix in the singular value decomposition in Eq. (2.44) give the empirical modes known as the proper orthogonal modes and the squares of the diagonal elements of Σ describe how much signal energy is captures by each mode. When used this way, the Karhunen-Loève expansion is similar to the modal analysis technique described in Section 2.2.4 in the approach to obtaining reduced models; namely, by assuming that the total response is a combination of some modal responses and retaining the dominant modes in the reduced model. Note, however, that the modes in the Karhunen-Loève expansion are empirical.

2.3.2. Balanced Truncation

The Karhunen-Loève expansion can be applied to a wide variety of dynamic system models for the purpose of modeling them properly. This includes linear and nonlinear, time-invariant and time-varying systems. The Karhunen-Loève expansion can also be applied to the same system for different state and input trajectories. This could ostensibly furnish significantly different proper models, each being "proper" only in the context of the trajectory used for obtaining it.

Balanced truncation is a special model reduction technique that involves applying the Karhunen-Loève expansion in particular ways to particular classes of systems. Its simplest rendition was originally proposed by Moore [141]. Specifically, Moore suggested the application of the Karhunen-Loève expansion to the state trajectory of the *balanced realization* of a linear and time-invariant system subjected to a series of impulses. A system's realization is balanced if its observability and controllability Grammians are equal, meaning that each state is as observable as it is controllable. When this is done, one finds that the less observable and less controllable states can be eliminated from the given system's model to furnish a reduced model. This *balanced truncation* process is a very interesting and powerful generalization of the Kalman canonical decomposition, which only eliminates the completely unobservable and completely uncontrollable states from a given system model to furnish a minimal realization of the model [142]. Note, however, that due to balancing the realization of the system changes, and balanced truncation is therefore not realization-preserving.

The balanced truncation technique proceeds mathematically as follows. First, it applies a state transformation to put the original model in a form where each state is equally controllable and observable. In this case, the controllability and observability matrices P and Q become diagonal, with the diagonal elements being the Hankel singular values, i.e., $P = Q = diag(\sigma_1, \sigma_2, \ldots, \sigma_n)$, where $\sigma_i = \sqrt{\lambda_i(PQ)}$ are the Hankel singular

values, which give a measure for the controllability and observability of corresponding states. Based on this measure, less controllable and observable states are truncated. There exists a global $\mathcal{H}_\infty$ error bound, which is the same as the $\mathcal{H}_\infty$ error bound in the Hankel norm approximation technique for the case when the $\tilde{D}$ matrix is not optimized, i.e.,

$$\left\| G - G_r \right\|_\infty \leq 2 \sum_{i=k+1}^{n} \sigma_i \tag{2.49}$$

where σ_i are the Hankel singular values of G corresponding to the truncated states. Note, however, that in Hankel norm approximation $\tilde{D}$ can be chosen such that only half of the $\mathcal{H}_\infty$ error bound of balanced truncation is achieved.

It is important to note the norm that is used in Eq. (2.49), because the singular values may not be as informative for other norms. As first shown by Kabamba, the singular values by themselves are not descriptive enough for the $\mathcal{L}_2$-norm of error [143]. Therefore, Kabamba introduced other invariants of the system, the *balanced gains*, that together with the singular values describe the contribution of each state to the $\mathcal{L}_2$-norm of the impulse response [143].

There is an interesting relationship between balanced truncation and singular perturbation. The *generalized singular perturbation approximation* allows for matching the magnitude of the original model at a desired frequency $s = s_0$, and choosing $s_0 = 0$ corresponds to the singular perturbation as given in Section 2.2.2, whereas choosing $s_0 = \infty$ corresponds to direct truncation [58]. Thus, assuming the original model is balanced, choosing $s_0 = \infty$ corresponds to balanced truncation, and furthermore, singular perturbation, i.e. choosing $s_0 = 0$, achieves the same error bound as the balanced truncation [58].

The literature describes many extensions of the above balanced truncation technique. These extensions include approximate balancing techniques that can be quite valuable when exact balancing is computationally costly [144-146]. Further extensions extend balanced truncation specifically to stochastic [147-149], passive [147], and

bounded real systems [150]. The literature also describes LQG balancing techniques for reduced order controller design [151] and frequency-weighted balanced truncation for reducing the approximation error over a specified frequency range rather than the whole spectrum [152-156]. Significant research has also pursued the balanced truncation of nonlinear systems [157-168]. This literature highlights the importance of balanced truncation, both as a powerful model reduction technique and as the basis for very extensive ongoing research, both theoretical and applied.

2.3.3. Component Cost Analysis

Another method that can be reviewed under the projection-based techniques category is component cost analysis [169-173]. In this approach, a cost function for the linear stable system

$$\begin{aligned} \dot{x} &= Ax + Bu, \quad x \in \mathbb{R}^n, u \in \mathbb{R}^m \\ y &= Cx, \qquad\quad y \in \mathbb{R}^p \end{aligned} \tag{2.50}$$

is defined as

$$V \triangleq \lim_{t\to\infty} E[\|y\|^2]; \quad \|y\|^2 = y^T y \tag{2.51}$$

This cost function satisfies the cost decomposition property

$$V = \sum_{i=1}^{n} V_i \tag{2.52}$$

where V_i is the contribution of the ith state, x_i, to the system cost, and is given by

$$V_i = \left[XC^T C \right]_{ii} \tag{2.53}$$

where X is the controllability Grammian, satisfying

$$XA^T + AX + BB^T = 0 \tag{2.54}$$

The reduced model is then obtained by truncating the low-cost states based on the rationale that the system cost should be perturbed minimally. However, it is important to know that deleting x_i, in general, does not necessarily cause a change of ΔV_i in V.

Note that the component cost analysis in this most basic form does not require a state transformation. Nevertheless, if the system is transformed into cost-decoupled coordinates, where XC^TC is diagonal, the component costs also quantify the amount by which the system cost will change if the corresponding states were truncated from the model. Furthermore, in these coordinates $n-r_C$ components will have zero component costs, where r_C is the rank of the matrix C. Therefore, in these coordinates a reduced model can be obtained that preserves the system cost. Cost decoupled coordinates are not unique, and one possible transformation into the cost-decoupled coordinates is given by

$$x = Tz$$
$$T = T_1 U;\ X = T_1 T_1^T;\ T_1^T C^T C T_1 = U \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} U^T \tag{2.55}$$

There is a close connection between component cost analysis and the idea of balanced gains introduced by Kabamba [143]. Specifically, if component cost analysis is applied to the balanced coordinates, the component costs exactly match Kabamba's results [171].

Furthermore, a very interesting relationship exists between balanced realization and cost-decoupled coordinates [173]. A generalization of the basic component cost analysis defines

$$\begin{aligned} y' \triangleq \begin{bmatrix} y \\ \dot{y} \\ \ddot{y} \\ \vdots \\ y^{(q-1)} \end{bmatrix} &= \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{q-1} \end{bmatrix} x + \begin{bmatrix} 0 & 0 & \cdots & 0 \\ CB & 0 & & \vdots \\ CAB & CB & & \vdots \\ \vdots & \ddots & & \vdots \\ CA^{q-2}B & \cdots & CB & 0 \end{bmatrix} \begin{bmatrix} u \\ \dot{u} \\ \ddot{u} \\ \vdots \\ \vdots \end{bmatrix} \\ &= \tilde{C}x + \tilde{D}u' \end{aligned} \tag{2.56}$$

and considers the system

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y' &= \tilde{C}x + \tilde{D}u' \end{aligned} \tag{2.57}$$

with the cost function

$$V \triangleq \sum_{k=1}^{m} \int_0^\infty \bar{y}^T(k,t) Q \bar{y}(k,t) dt \tag{2.58}$$

where $\bar{y}(k,t)$ is the response of the system for an impulse at the k^{th} input channel while all other inputs being zero, and Q is a weight matrix. Then, the cost-decoupling transformation

$$\tilde{T} = T_1 U \begin{bmatrix} \Sigma^{-1/4} & 0 \\ 0 & I \end{bmatrix} \tag{2.59}$$

yields the balanced coordinates, if

$$q = n,\ Q = \int_0^\infty \begin{bmatrix} \alpha_0(t) \\ \vdots \\ \alpha_{q-1}(t) \end{bmatrix} I_p \begin{bmatrix} \alpha_0(t)I & \cdots & \alpha_{q-1}(t)I \end{bmatrix} dt \tag{2.60}$$

where $e^{At} = \sum_{i=0}^{n-1} \alpha_i(t) A^i$. These results imply that balanced coordinates are a special case of the generalized cost-decoupled coordinates, and thus the component cost analysis is a generalization of the balanced truncation.

2.4. OPTIMIZATION-BASED TECHNIQUES

The frameworks of both frequency- and projection-based proper modeling techniques are based on the same goal: to identify and isolate the dominant characteristics of a given model. For frequency-based methods these characteristics lie in the frequency domain, and for projection-based methods they are in the state space.

In addition to this rather intuitive and practical motivation of retaining the model's dominant characteristics, one may also seek to formally achieve a minimal difference between the predictions of the full and reduced models subject to a complexity constraint. Such techniques are referred to as *optimization-based* proper modeling techniques in this work.

Optimal Hankel norm approximation, for instance, is an optimization-based proper modeling technique, because it seeks to minimize the Hankel norm of the difference between a full model and a reduced model, subject to a bound on the reduced model's order. The fact that optimal Hankel norm approximation is also a frequency-

based proper modeling technique underscores the fact that the classification of proper modeling techniques used here, while intuitively appealing, is certainly not strict. Interestingly, optimal Hankel norm approximation is also a projection-based proper modeling technique. This raises an important question, namely, whether one can formulate a "unified" model reduction problem: one that simultaneously seeks optimality in the frequency and state space domains.

The above question was partly answered by Hyland and Bernstein's seminal work on the *optimal projection equations* [174]. In this work, Hyland and Bernstein formulated the proper modeling problem as a problem of minimizing a quadratic measure of the error between a full model and its proper counterpart, subject to implicit rank constraints on the proper counterpart. This furnished a set of first-order necessary conditions for optimality of the reduced proper model, which Hyland and Bernstein expressed as a coupled system of two Lyapunov equations. Hyland and Bernstein then studied balanced truncation in the context of these necessary conditions for proper model optimality. They found that balanced truncation furnished reduced models that deviate significantly from quadratic optimality: a conclusion also supported by earlier research by Kabamba [143]. The significance of this finding cannot be overemphasized. It highlights the fact that a "proper" model developed using one metric (e.g., the relative observability and controllability of different states) can be far from being "proper" in the context of a different metric (e.g., quadratic optimality). In other words, there is no universal proper modeling algorithm applicable to all systems under all circumstances. Rather, different proper modeling algorithms are better suited to different problems, and one should carefully select the proper modeling metric ideally suited for the problem at hand.

Optimization-based proper modeling techniques typically seek to minimize the $\mathcal{L}_2$, $\mathcal{H}_2$, or $\mathcal{H}_\infty$ norm of the difference between a given "full" model and its proper counterpart, subject to a constraint on the order (i.e., "complexity") of the proper counterpart. Wilson, for instance, was the first to solve the $\mathcal{L}_2$ optimal reduction problem

[175]. Specifically, he minimized the integral square error of the difference between the impulse responses of the full and reduced models. Other examples employing the $\mathcal{L}_2$ norm in their cost function include [176-186]. Similarly, Luus optimizes a reduced model to minimize the deviation of its frequency response from that of the corresponding full model [187]. As another example, Sou proposes a relaxation to the non-convex optimal $\mathcal{H}_\infty$ norm reduction problem to turn it into a quasi-convex problem, which is easier to solve due to the existence of a unique solution [188]. The proper modeling problems resulting from such formulations often do not have analytic solutions, and must hence be solved numerically.

As a result, much of the optimization-based proper modeling literature focuses on the development of numerically efficient optimization algorithms, with special attention to the convergence properties of these algorithms. Gouda *et al.*, for instance, obtain a proper model of a building's thermal response using sequential quadratic programming [189]. Similarly, Hachtel *et al.* propose an interactive optimization technique incorporating linear programming as a tool for nonlinear model reduction [190].

Both linear and sequential quadratic programming are local search techniques that may not be able to find globally optimal proper models. With this in mind, Assunção and Peres propose a branch-and-bound algorithm for the solution of the optimal $\mathcal{H}_2$-norm-based proper modeling problem [191]. Finally, Chen and Fang [192], Spanos *et al*. [193], and Ferrante *et al.* [194] propose reduced model optimization algorithms that have attractive mathematical guarantees of convergence.

Optimization-based approaches may or may not be realization-preserving, depending on whether they fix the given system's realization during the search for an optimal reduced model or allow it to vary. While most optimization-based approaches in the literature are not realization-preserving, it is certainly possible to construct ones that are.

2.5. ENERGY-BASED TECHNIQUES

Energy-based proper modeling techniques are built on the intuitive fundamental premise that in an energetic system, the most important components to model accurately are those characterized by the largest magnitudes of energy (or power) flow. Therefore, these algorithms reduce a given model by eliminating less energetic components, while trying to minimize the effect of the elimination on the overall energy flow. The well-known Rayleigh-Ritz method exemplifies this perspective on model reduction [68]. Methods that utilize the $\mathcal{H}_2$-norm (e.g., [191]) or the $\mathcal{L}_2$-norm (e.g., [175]) can also be classified as energy-based. Other energy-based model reduction algorithms include statistical energy analysis [195] and the power-based model reduction algorithms by Rosenberg and Zhou [196, 197].

Rosenberg and Zhou's model reduction algorithm [196, 197] is based on the intuitive notion that in an energetic dynamic system those components characterized by higher mean-square energies should be more important to model than those characterized by lower mean-square energies. This leads to a simple, intuitive, realization-preserving, and powerful model reduction technique with no theoretical proof for convergence, reduced model stability, or "optimality".

Louca *et al.* extend Rosenberg and Zhou's algorithm by proposing a new energy-based model reduction metric called *activity* [20]. The activity of an energetic element is defined as the time integral of the absolute value of the power flowing through it over a particular time-window for a particular input. In a bond-graph setting, where the flow through an element *i* and the effort across it are denoted as f_i and e_i, respectively, the element's activity is defined as

$$A_i = \int_0^T \left| f_i^T(t) e_i(t) \right| dt \tag{2.61}$$

where *T* is the width of the desired time-window. The activity of an element can, hence, be physically interpreted as the total energy flow through the element within a specified

time-window for a specific input. It can also be interpreted as the $\mathcal{L}_1$-norm of the power flow through the element, multiplied by the width of the time window used to compute that norm.

Louca *et al.* conjectured that in an energetic system, the more active elements are more important to model than the less active elements. An *element*, in this context, is any component in the system's bond-graph representation, including generalized resistors, capacitors, and inductors. Based on this conjecture, Louca *et al.* proposed an activity-based realization-preserving *model order reduction algorithm (MORA)* [20], and developed techniques for physically interpreting the reduced models generated by this algorithm [198].

The fundamental premise behind MORA, namely, that activity can be used as a proper modeling metric, is mostly intuitive. However, it is supported by some important application studies [199-201]. Furthermore, recent work by Fathy and Stein has unveiled *fundamental concordances* between MORA and balanced truncation [202]. These concordances are special cases where the two algorithms are mathematically guaranteed to furnish identical reduced models. While these concordances do not provide a general mathematical foundation for MORA, they do lend credence to MORA as a mathematical model reduction algorithm, at least in the special cases covered by the concordances [202].

Beyond its viability as a model reduction metric, activity has also proven viable as a model *partitioning* metric. Specifically, Rideout *et al.* use activity to quantitatively and systematically look for decoupling among the elements of a model and to partition the model based on the discovered decoupling [19]. Once the partitions are obtained, the simulation can be carried out either by simulating the driving partition first and using its output as an input to the driven system, or, in case only the driving partition is of interest, by completely eliminating the driven partition and keeping only the driving partition.

The methods discussed above are based on some norm of power as a metric. Ye and Youcef-Toumi propose an alternative to this approach [203]. Specifically, they propose a sensitivity analysis that is applied to the bond energies in a bond graph model. Their analytical approach is limited to systems in which the energy levels of independent energy-storing components are piecewise invertible functions of their corresponding states. However, due to the sensitivity-analysis-based framework of the approach, better results may be obtained compared to the metrics above, though at the cost of ease of implementation.

The works discussed above do not provide bounds on the effects of the neglected dynamics on the reduced model. Knowledge about such bounds is important especially if the reduced model is going to be used for controller design. Specifically, such bounds are important for the robustness of the controller. It is known that the closeness of the reduced model to the full model in terms of an energy-based metric does not guarantee that the stable controller designed for the reduced model will also be stable for the full model. In fact, the weakest norm that can guarantee this stability is the gap metric [204-208]. To address this issue, Chang et al. take a Lyapunov function approach to energy-based model reduction [209]. This allows them to provide bounds on the disturbances to the reduced model caused by the unmodeled dynamics. Even though implementation-wise this method is not as easy as, e.g., the activity metric, the additional information about the bounds can be invaluable for robust controller design.

2.6. DISCUSSION AND CONCLUSIONS

The process of modeling a dynamic system invariably entails a tradeoff between model accuracy and simplicity. Simpler models can be easier to simulate, analyze, comprehend, and control than more complex ones, but this often comes at the expense of

accuracy and, hence, potential viability. Recognizing this fundamental tradeoff, the literature deems a model "proper" if it balances the needs for accuracy and simplicity.

The formal definition of "proper" models may be relatively new [59], but its underlying emphasis on the need for balancing model fidelity and complexity has been recognized for many decades. In fact, the literature presents many techniques for *reducing* complex models until they become proper, or *deducing* proper system models from simpler subsystem models.

This chapter briefly surveys these techniques and classifies them into *frequency-*, *projection-*, *optimization-*, and *energy-based* depending on their underlying metrics for assessing the relative importance of a model's different dynamics and subsystems. This classification is neither well-established nor strict, as evident from the fact that a given proper modeling algorithm often belongs to more than one of these categories. However, it is adopted herein for its convenience for presentation.

A careful examination of the different proper modeling techniques in the literature leads to the fundamentally important conclusion that *there is no universal proper modeling technique suitable for all modeling problems and all applications*. Rather, different proper modeling techniques are often better suited to different problem spaces.

Despite the richness of the proper modeling literature, many important problems remain to be addressed. In particular, one may claim that each of these techniques has one or more of the following limitations:

1. Applicability to a limited set of systems: Some proper modeling techniques are limited to particular classes of systems. For example, polynomial approximation methods are applicable to linear systems only, and singular perturbation is ideally suited for systems with multiple time scales.

2. Requiring a realization change: Many proper modeling techniques (e.g., balanced truncation, optimal Hankel norm approximation, proper orthogonal

decomposition, etc.) project the dynamics of the given system onto a new state space conducive to reduction. Although this may yield better results in terms of minimizing the approximation error, one may also wish to preserve the original realization, perhaps because of its intuitive appeal.

3. Trajectory independence: Many proper modeling techniques seek models whose accuracy is acceptable over a broad range of state and input trajectories. One example is $\mathcal{H}_\infty$ model reduction, which seeks to reduce a given model while minimizing the resulting error over the entire frequency spectrum. Such trajectory independence is often attractive, but one may conceivably seek a model that is proper only for a given trajectory or small family of trajectories. In such situations, trajectory-independent proper modeling algorithms may furnish excessively complex models, and trajectory-dependent approaches may be preferable.

4. Being limited to equation level: Graph representations (e.g., linear graphs, bond graphs, etc.) often provide intuitively appealing depictions of system models, but most proper modeling algorithms operate at the equation level. This means that even though the equations derived from a graph representation could be reduced using the existing techniques, the reduction would not necessarily be reflected at the graph level, which may hinder the advantages of having a graph-level representation.

5. Not considering the structure of the model: Most proper modeling methods seek to reduce the *order* of a given model, i.e., its number of states. Therefore, possible reductions in the *structure* of the model, i.e., how the components of the given system interact with each other, are typically not taken into account.

For brevity, the chapter focuses mostly on the deterministic proper modeling problem. The notion of a "proper model" becomes particularly powerful in the context of systems with significant uncertainties. In particular, when modeling a stochastic system, one may legitimately ask: which of the system's various uncertainties are more important to model, and which are negligible? This leads to the notion of a *stochastic proper model*: one capturing only the most salient dynamics *and uncertainties* of a given system. Significant research exists, and continues, in the area of stochastic proper modeling, but this chapter focuses on deterministic proper modeling for brevity.

Finally, it is important to note that proper models of dynamic systems are often a means to an important practical end. In particular, the ultimate goal of any proper system modeling exercise is often to not only better understand the system's behavior, but also to use this understanding as a means towards better system designs and controls. This implies that a proper model must, therefore, be both *scalable* and *control-oriented*. A system model is *scalable* if it captures not only the dynamics of a given system, but also how these dynamics change with system design parameters. Furthermore, a system model is *control-oriented* if it accurately captures those dynamics that are most important for the effective control of the given system. Both scalable and control-oriented modeling are rapidly becoming active research topics, and a thorough discussion of these topics is omitted from this chapter for brevity.

CHAPTER 3

STRUCTURAL SIMPLIFICATION OF MODULAR BOND-GRAPH MODELS BASED ON JUNCTION INACTIVITY

3.1. INTRODUCTION

The modular modeling paradigm facilitates the efficient building, verification and handling of complex system models by assembling them from general-purpose component models. A drawback of this paradigm, however, is that the assembled system models may have excessively complex structures for certain purposes due to the amount of detail of the component models, which has been introduced to promote modularity. For example, a multibody system can be modeled using generic rigid-body models with 6 degrees-of-freedom (DoF) to represent the components of the system, but then constraints have to be added to the model to match the actual DoF of the system.

This chapter presents a domain-independent structural simplification technique that can detect such unnecessary complexities in a modular bond-graph model and eliminate them from the model without compromising accuracy. To this end, the activity concept in the literature is extended to define "inactivity" for junction elements, and simplification is obtained by detecting and eliminating inactive junction elements and by propagating the implications. It is shown by example that this simple idea can result in models that are conceptually and computationally more efficient than the original modular models. The realization-preserving and input-dependent characteristic of this approach is highlighted.

3.2. MOTIVATION

One possible approach to modeling dynamic systems is to develop modular models for the components first and then to assemble the system model by combining the component models in accordance with the system topology [210]. Such a modular approach is well established in the multibody dynamics area, for example, where the system model is obtained by augmenting 6-DoF rigid-body models with constraints. Commercial software based on this idea is available (e.g., ADAMS [211]).

The modular approach has many well-established advantages. These include independent creation and reuse of submodels, hierarchical model structure, ease of adjustment of model complexity, ease of model verification and ease of handling large systems.

A drawback of this approach, however, is that when general-purpose component models are assembled into a system model, the resulting model can be excessively complex for a given application [9]. The component models need to be created for a broad range of applications, and therefore need to include a lot of detail relevant for that scope. In the case of modular modeling of multibody systems, for example, a generic component model for a rigid body may consider all possible motions in space and include all 6 DoF. However, when component models are assembled into a particular system model, some of that detail may become irrelevant/unimportant in that particular context. Returning to the multibody example, the model for a particular system can be obtained by augmenting component models of rigid bodies with relevant constraints, but then the number of DoF of the system is less than the sum of the number of DoF of the unconstrained components. Therefore, the system model includes an excessive amount of complexity, and it may be desired to eliminate this excessive complexity.

One motivation for eliminating excessive complexity could be that a simpler model could prove more insightful by showing only what is of relevance to the problem

at hand. Also, a simpler model would typically have fewer parameters and states, which reduces the number of parameter values and initial conditions that need to be identified. Note that an initial estimate of parameters is necessary for the full model, but a costly accurate identification can be delayed until the simplified model is obtained. If the model is going to be involved in a control design problem, a simpler model could also simplify the control problem [69]. Finally, a simpler model is generally more computationally efficient, which makes the model more suitable for iterative processes, such as optimization [212], sensitivity analysis [213, 214], Monte Carlo simulation [215], system identification [216], etc., or for real-time simulation [217]. It is acknowledged, however, that simplicity does not necessarily always imply computational efficiency [211].

As mentioned in Chapter 1, the complexity of a model can be reduced in different ways. Specifically, it was distinguished between partitioning and simplification as special cases of reduction. This chapter focuses exclusively on simplification.

It was also mentioned in Chapter 1 that as a way of expressing the models in a modular modeling environment, bond graphs prove to be suitable due to their power-based graphical nature and their lending themselves to modularity [21]. As an example, a bond-graph-based modular modeling paradigm has been shown to be suitable for modeling reconfigurable machine tool servo drives [9]. There are other possible representations to create a modular environment, such as block diagrams [210] or Modelica [218], but bond graphs are preferred in this work due their convenience. Nevertheless, it is noteworthy that although the method presented in this chapter is particularly amenable to bond graphs, it is applicable to other representations as well.

The literature presents many mathematical tools, such as pole-zero cancellation [13], Kalman's minimal realization [14], or explicit elimination of Lagrange multipliers [15, 16], that can be used for model simplification, but these techniques apply at the equation level. Thus, even though the equations derived from a bond graph could be simplified with such techniques, the bond graph itself would still contain the excessive

complexities, obscuring insight, and requiring the repetition of the equation-level simplification each time the equations are derived from the bond graph.

There are some well-established rules for bond-graph-level simplification [21], such as eliminating loose or power-through junctions, merging adjacent junctions of same type, eliminating a null effort (flow) source connected to a 1-junction (0-junction), lumping dependent elements, or some structural equivalencies. However, these rules by themselves are not enough to eliminate a Lagrange multiplier from a bond graph, for example. Hence, they still leave room for more simplification.

A more advanced model simplification procedure in the bond-graph domain is proposed by Rinderle and Subramaniam [219], consisting of four main steps: (1) eliminating null sources, (2) eliminating transformers with zero modulus, (3) eliminating constraining junction structures, and (4) reducing the number of dependent inertia elements through parameter lumping based on the Lagrangian. Although very useful in some cases, this technique is not very effective when the transformers are modulated, or the constraints are enforced through Lagrange multipliers, for example, so that there are no constraining junction structures in the model.

There are also some bond-graph level tools that have been developed for reduction and partitioning purposes, but can serve, to some degree, for simplification purposes as well. In particular, as reviewed in Chapter 2, a metric called "activity" for measuring power flow among bond-graph elements has been proposed by Louca *et al*. and used to create a model reduction algorithm called MORA [20]. If MORA is used to remove zero-activity elements only, it serves as a simplification tool. Nevertheless, a significant amount of complexity may still remain in the model, because MORA concentrates only on the energy elements and not on the junction structure. In addition, the partitioning algorithm proposed by Rideout *et al*. also leverages the activity concept [19]. Weak coupling points can serve for simplification purposes if the driven partition consists only of an inactive junction structure and elements that can be removed along

with the junction structure, but this is not always the case. Thus, this method is also not very effective for simplification.

This chapter presents a realization-preserving input-dependent algorithm to achieve structural simplification of nonlinear models at bond-graph level. To this end, a unified treatment of junction elements is introduced. In particular, the activity metric [20] is extended to junction elements to identify the inactive ones and eliminate them to simplify models. This approach can be considered as a generalization of the well-known idea that 1-junctions with zero flow and 0-junctions with zero effort can be eliminated from a bond graph. Thus, the contributions of this chapter are the generalization of the zero-flow and zero-effort metrics into a unified metric, inactivity; a procedure to identify and remove inactive junction structures; and a detailed discussion of the realization-preserving and input-dependent property of this approach.

3.3. MODEL SIMPLIFICATION BASED ON JUNCTION INACTIVITY

3.3.1. Inactivity of Junction Elements

The activity metric developed by Louca *et al*. is a measure of power flow in a model for a given input [20]. Activity of an energy element (i.e., generalized inertia (I), capacitance (C) or resistance (R)) is formally defined as

$$A = \int_{t_1}^{t_2} |e \cdot f| \, dt \tag{3.1}$$

where A, e and f are the activity, generalized effort and flow of the element, respectively. Based on the hypothesis that elements with low activity contribute less to the system dynamics, the activity metric is used as the basis for a systematic model reduction technique called “Model Order Reduction Algorithm (MORA)” [20]. In MORA, the activity metric is defined and used for assessing the single port energy elements (I, C and R) only.

If the junction structure is also to be considered for simplification, a metric for junctions is also necessary. In this work a junction element, 1- or 0-junction, is called "inactive," if all the bonds that are connected to the junction element have a negligible activity, i.e.,

$$A_i = \int_{t_1}^{t_2} \left| e_i f_i \right| dt < \varepsilon, \quad i = 1, 2, \ldots, n \tag{3.2}$$

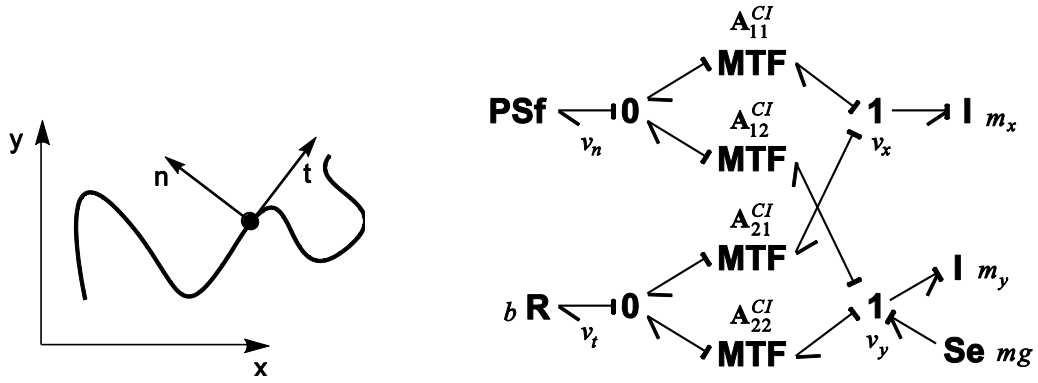
where A_i, e_i and f_i are the activity, effort and flow of bond *i* connected to the junction element, respectively; ε is a small number, which will be referred to hereafter as the inactivity threshold; and *n* is the number of bonds connected to the junction element.

The hypothesis is that inactive junction structures can be removed from the bond graph without compromising accuracy, thus simplifying the model.

Two things are important to note here: First, the inactive junction concept can be considered as the generalization of the idea that 1-junctions with zero flow and 0-junctions with zero effort can be eliminated from a bond graph without sacrificing the accuracy of the model. This is because a 1-junction (0-junction) will be inactive not only if its flow (effort) is zero, but also if the efforts (flows) are zero. For example, if a 1-junction represents a characteristic non-zero velocity component along which no force does any work, then that 1-junction is going to be inactive despite the non-zero flow.

Second, the elimination of an inactive junction does not necessarily correspond to removing every null-power bond from the model. In fact, the latter may lead to computational problems. To illustrate this, consider a particle of mass *m* that is constrained to move along an arbitrary path (Fig. 3.1a). The particle experiences a gravitational force in the –y direction, and a viscous friction, *b*, as it moves along the path. Assume that the modeler has chosen to work with two coordinate frames, the inertial (x-y) and constraint (n-t) frames, for their convenience to express the gravitational and constraint forces, respectively. The bond graph of this scenario is given in Fig. 3.1b, where the pseudo-flow source (PSf) [9] represents the Lagrange multiplier

enforcing the constraint, and modulated transformers take care of the coordinate transformation between the inertial and constraint frames. Specifically, $[v_n \quad v_t]^T = \mathbf{A}^{CI}[v_x \quad v_y]^T$, where $\mathbf{A}^{CI}$ is the transformation matrix from the inertial frame to the constraint frame. Note that the elements of $\mathbf{A}^{CI}$ can become arbitrarily small, since the path is arbitrary. In other words, even though $\mathbf{A}^{CI}$ will always be invertible, its individual elements may be not. Therefore, all the modulated transformers in Fig. 3.1b should have a fixed flow-in-flow-out causality to avoid singularities. In that case, the only possible causal assignment is as shown in Fig. 3.1b. Note that the bond between the PSf and 0-junction elements will have null power due to the zero normal component of the particle velocity. However, it cannot be removed from the bond graph, because that would create a causal conflict due to the fixed causal assignment of the MTF elements. The proposed metric would identify the 0-junction to which the PSf element is connected to as active and would not consider it for simplification, thereby keeping a null-power bond in the model.



(a) (b)

Figure 3.1. (a) Particle constrained to an arbitrary path, (b) bond-graph model of the system

3.3.2. The Simplification Procedure

Using the inactivity metric defined in the previous subsection for a junction element, a procedure to simplify bond-graph models is proposed below.

Note, because activity depends on the excitation and the time-window selected for the activity calculation, it is critical to choose them properly for the task at hand. Selecting proper excitations and the time-window are considered as prerequisite steps. Section 3.4.2 will discuss the importance of proper excitations more in detail.

After the prerequisites are satisfied, the main steps of the simplification procedure can be described as follows:

1. Detection of inactive junction elements: A simulation run is performed and the activity values are recorded. The inactivity threshold ε is selected, typically on the order of magnitude of the numerical integration tolerance, and the inactive junctions are identified using Eq. (3.2).

2. Preserving modulating signals: Some junction structures may be inactive, yet important for generating a modulating signal necessary for the rest of the bond graph. To preserve the modulating signal, the inactive bond-graph junction structure that generates the modulating signal should be converted into a block diagram, instead of being removed completely.

3. Elimination of inactive junctions: The remaining inactive junctions are removed from the model, along with the elements and submodel ports that become detached after the removal of the inactive junction elements.

After the last step, if desired, the well-known bond-graph simplification techniques can be applied, such as removal of power through 1- or 0-junctions, removal of unity gain transformers, merging adjacent junctions of the same type, or lumping dependent inertias onto independent ones.

This procedure is given as a flowchart in Fig. 3.2 and is illustrated with an example in Section 3.3.3.

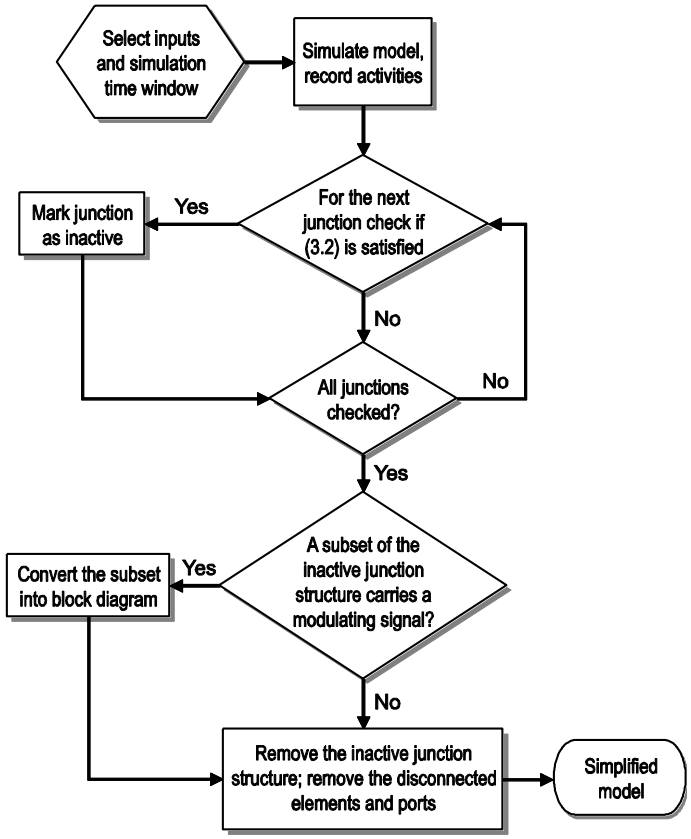
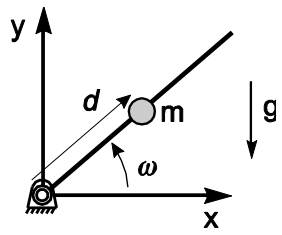


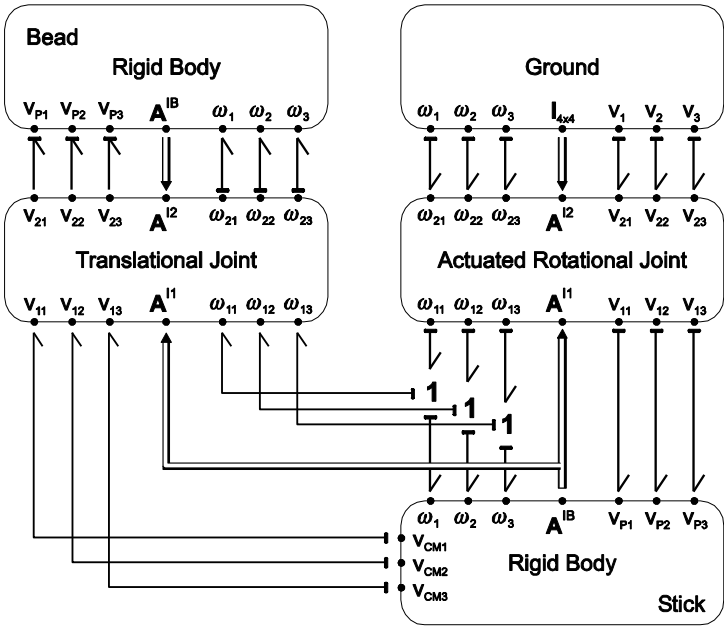
Figure 3.2. The flowchart of the simplification process

3.3.3. Example: Bead on a Stick

To illustrate how the proposed algorithm works, consider a bead that can slide smoothly on a stick, which is swung in a vertical plane with constant angular velocity ω. Figure 3.3 illustrates the described system and shows its modular bond graph, which is composed of 3D rigid-body and joint models. For the details of the modules please see the Appendix A. One may argue that this system is too simple to justify the modular modeling approach, but since the purpose here is to demonstrate and discuss the proposed



(a)



(b)

Figure 3.3. (a) The bead and stick system, (b) its modular bond-graph model

simplification algorithm, all the example systems in this chapter are intentionally kept simple.

Note that the modular model (Fig. 3.3b), although very easy to create, is quite complex considering the given 1-DoF system: The model has two 6-DoF rigid-body models representing the stick and bead, and two joint models representing the connections between the bead and stick, and the stick and ground. As a result, there are many dimensions, in which the system cannot move, but which are included in the model anyway because of the modular approach adopted in creating it. For the bond-graph this

implies that there are many unnecessary elements in the model. Thus, there is an opportunity for simplification created by the modular approach to modeling.

Simplification is then carried out as follows: A simulation run is performed to record the activity values, and the original model is simplified based on an inactivity threshold of 10^{-5}, which is also equal to the simulation tolerance. The resulting simplified model is given in Fig. 3.4. As seen in Fig. 3.4, the proposed algorithm significantly reduces the complexity of the bead and joint models, and completely removes the stick and ground models. A detailed explanation of the simplification and the physical motivation behind it is given next.

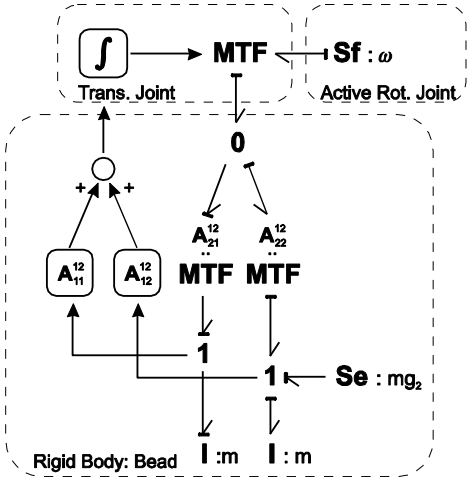


Figure 3.4. The simplified bead and stick model

First, consider the bead model. Figure 3.5 shows the original rigid-body module representing the bead and its simplification in detail. In Fig. 3.5 the inactive junction-structure to be removed is shown in grey, along with the elements to be removed based on the implication of the inactive junction-structure. The junction structure that is inactive, but important for the generation of a modulating signal, is shown with outlined characters and the corresponding bonds are denoted by dash-dotted lines.

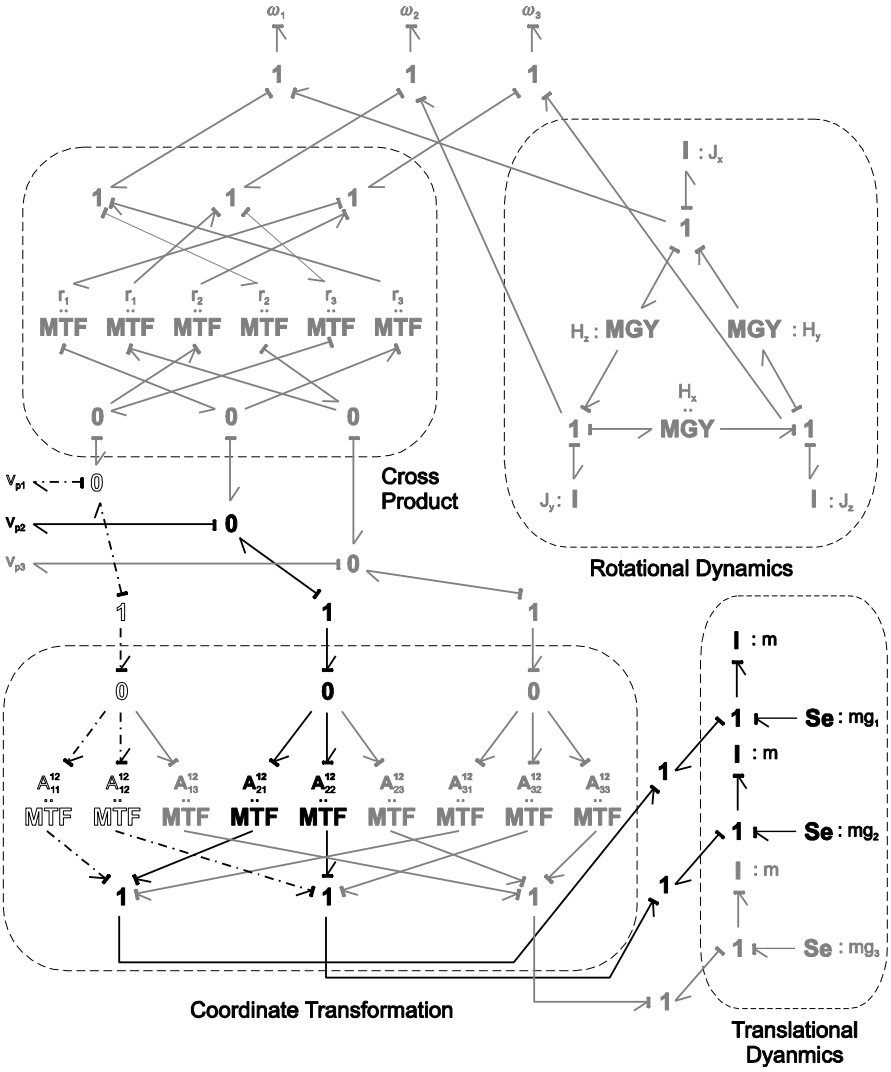


Figure 3.5. Bead model with inactive junction structure and its implications marked

The phenomena that are removed due to inactivity or preserved despite inactivity, and the physical motivations behind those removals and preservations are as follows: Rotational dynamics are eliminated completely, because the stick and, therefore, the bead connected to it are rotating with a constant angular velocity. The cross-product $\vec{\omega}\times\vec{r}$, i.e., the velocity of the connection point of the bead due to rotation, is removed, because the connection point coincides with the center of mass of the bead, i.e., $\vec{r}=0$. The translational dynamics in the z-direction is removed, because the system is planar. Due to the same reason, there is also some simplification in the coordinate transformation from

bead frame to the inertial frame. Finally, the preserved inactive structure is generating the signal $\dot{d}$, i.e., the velocity of the bead relative to the stick. The inactivity of this particular structure is due the smoothness of the motion. Nevertheless, the signal $\dot{d}$ is important, as its integral, d, i.e., the position of the bead relative to stick, is one of the states of the system. The associated force variable, however, is not necessary, because there is no force that acts along d. Hence, this particular junction structure is converted into a block diagram in Fig. 3.4, which is the signal path from the 1-junctions in the bead model to the integrator in the translational-joint model. This eliminates the corresponding force variable from the model.

Second, consider the stick and ground models. Note how the algorithm recognizes the fact that the dynamics of the stick are irrelevant due to the constant angular velocity, and eliminates the rigid-body model representing the stick completely from the model. This makes the ground model redundant as well.

Finally, with 4 DoF removed from the bead model and the stick and ground models eliminated completely, the constraints in the joint models become unnecessary and are therefore removed, which significantly simplifies the joint models as well.

Note that all these physical observations are incorporated into the model automatically by the simplification algorithm. Therefore, the simplification provides additional insight into the system, making the model more conceptually efficient.

The simplification of the bond graph leads not only to conceptual efficiency, but also to computational efficiency. With a reduced number of simulation equations, as evident from Table 3.1, the simplified model takes nearly 80% less time than the original model for 1000 iterations of simulation while predicting the same behavior as the original model. The rationale behind repeating the simulation 1000 times is to reduce the stochastic variations in simulation time due to the other processes concurrently running on the computer. The quantities in Table 3.1 are obtained with the 20-sim software [220].

Table 3.1. Increase in efficiency due to simplification

	Original model	Simplified model
Number of equations	557	37
Number of variables	698	53
Number of independent states	20	4
Number of dependent states	9	1
Number of constraints	2	0
CPU time for 1000 runs	17.33s	3.64s

3.4. DISCUSSION OF PROPERTIES OF INACTIVE-JUNCTION-BASED SIMPLIFICATION

The proposed algorithm has two important properties: being realization-preserving and input-dependent. It is realization-preserving, because the simplified bond graph is in essence a subset of the initial bond graph. This property is important to preserve the physical meaning of the original model, but has further implications as will be discussed further in this section. The proposed algorithm is input-dependent, because the power flow in the model and therefore the activity analysis depends on the inputs. Here the term "input" collectively refers to excitation, parameters, and initial conditions. This can be considered as both an advantage and disadvantage as will be demonstrated later in this section.

3.4.1. Preservation of Realization

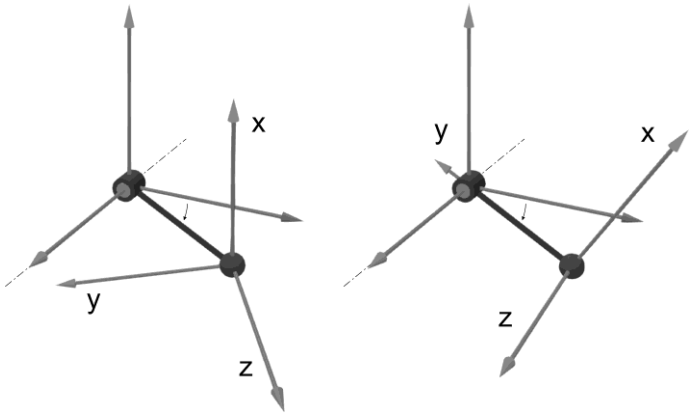
The proposed algorithm simplifies a given bond-graph model by detecting and eliminating the elements that do not contribute to the system dynamics. As such, the simplified model is a subset of the initial bond-graph. Therefore, the realization of the original model is preserved. This realization preserving property is important to preserve the physical meaning of the original model. As a result of this property, the realization of the original model significantly affects the outcome of the algorithm, because some realizations can be more conducive to simplification than others. This is demonstrated by two mechanical system examples below. In particular, the first example highlights the

effect of the orientation of the coordinate frames, and the second one highlights the effect of using absolute vs. relative coordinates.

Example 1: Orientation of the Coordinate Frames in Multibody Systems

In the mechanical domain the orientation of the coordinate systems is an important factor in simplification, because some coordinate systems can be more conducive to simplification than others. If the coordinate systems of the original model are not properly aligned with the motions and constraints, then performing a coordinate transformation first could yield more dramatic simplification results.

As an example, consider the simple pendulum shown in Fig. 3.6 with two possible choices of coordinates for this system, where the first one is rather arbitrary and the second one is aligned with the rotation axis and the pendulum itself. The original modular model for both choices of coordinates is given in Fig. 3.7. Due to poor alignment, the first frame (Fig. 3.6a) hardly yields a simplification, as shown in Fig. 3.8. If, however, the pendulum coordinate frame is oriented as in Fig. 3.6b, the junction structure corresponding to the rotation of the pendulum about the x- and y-axes becomes inactive, because the angular velocity of the pendulum is completely described by its z-component when resolved in this particular coordinate frame. In this case one obtains the simple model shown in Fig. 3.9.



(a) (b)

Figure 3.6. Two possible choices for pendulum coordinate frame: (a) arbitrary, (b) aligned with motion

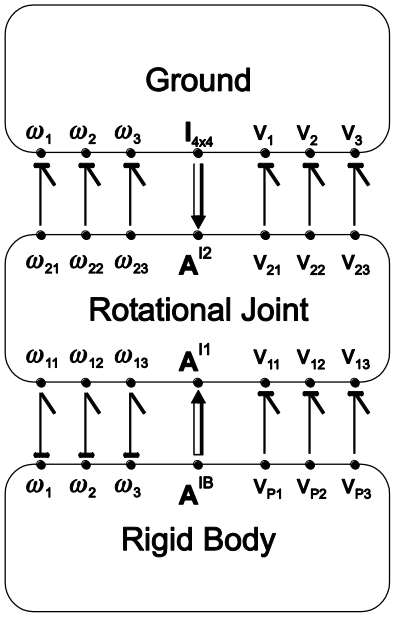


Figure 3.7. The modular pendulum model

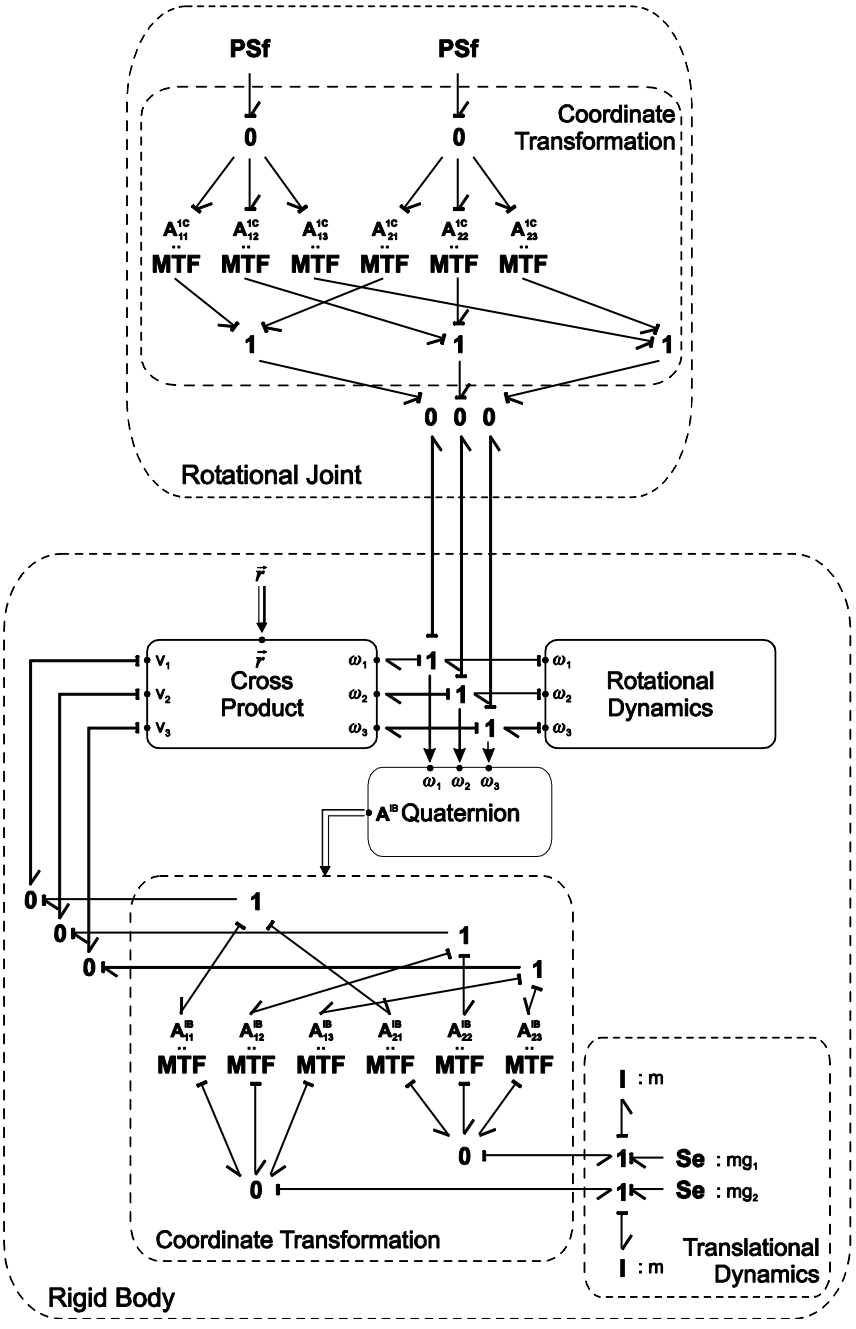


Figure 3.8. The simplified pendulum-model for the first coordinate frame

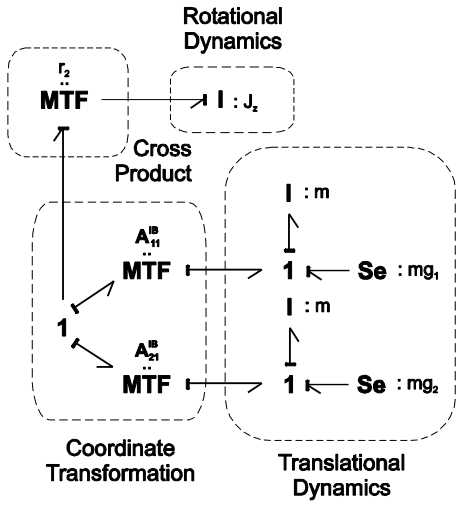


Figure 3.9. The simplified pendulum-model for the second coordinate frame

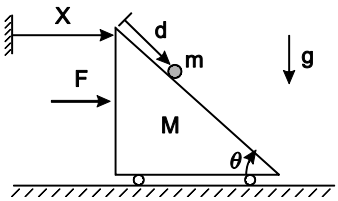
An automated way to switch from the first frame to the second one could be a desirable as a supplement to the proposed algorithm. Such a procedure is proposed in the next chapter that employs the Karhunen-Loève expansion to detect the existence of and to find the transformation into a better aligned coordinate frame.

Example 2: Choosing Between Absolute and Relative Coordinates

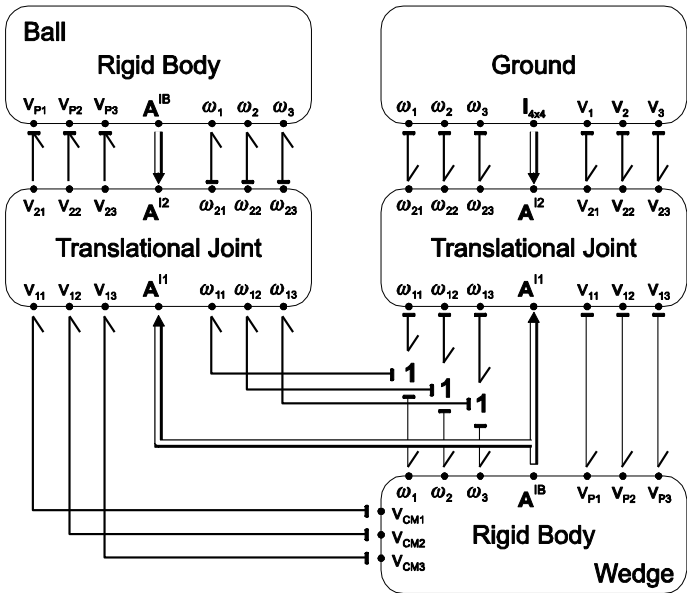
Reorienting the coordinate frames may not by itself furnish the realization most conducive to simplification. Another realization-related question to consider is whether to use absolute or relative coordinates. This example shows how the two different choices yield different simplification results.

Consider a ball, which can move smoothly on the oblique surface of a wedge, as shown in Fig. 3.10a. A horizontal force acts on the wedge in the direction shown, so that the system starts moving from rest. Figure 3.10b shows the corresponding modular bond graph.

The bond graph as given in Fig. 3.10b uses absolute coordinates to express the dynamics, i.e., all the inertia elements are connected to 1-junctions that represent an absolute velocity, whether it is resolved in the inertial frame or a body coordinate frame.



(a)



(b)

Figure 3.10. (a) The ball and wedge system, (b) its modular bond-graph model

In this case, no matter how the coordinate frames are oriented, the simplification procedure will always end up with two translational degrees of freedom for the ball along with a constraint, because the ball moves in two dimensions with respect to the ground.

It is easy to see, however, that if the generalized coordinates are chosen as X and d (Fig. 3.10a), it is possible to express the system dynamics in only two second-order differential equations with no constraints. To get the same result with the simplification procedure, the same coordinates must be used in the bond graph, i.e., the translational dynamics of the ball must be expressed using relative coordinates, since the second generalized coordinate is a relative quantity. In bond-graph terms, the inertia elements

that are currently connected to the 1-junctions representing the ball's absolute velocity in the inertial coordinate frame must be transferred over to the 1-junctions that stand for the ball's relative velocity with respect to the wedge and reside in the translational joint block. This transformation is straightforward and hence its details are omitted for brevity.

When the aforementioned modification is made to the system bond graph and the simplification procedure is applied, one obtains the result shown in Fig. 3.11, where there are only two degrees of freedom and no constraints.

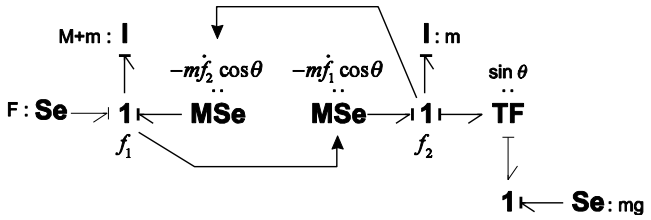


Figure 3.11. Simplified ball and wedge model when relative coordinates are used

This example shows that if the coordinates used to express the dynamics are independent, as is the case when relative coordinates are used, the proposed algorithm will eliminate all the constraints from the model. Although not essential for the proposed algorithm to work, an automated way to convert a realization in absolute coordinates into a realization in relative coordinates could be another beneficial supplemental tool for the proposed algorithm.

3.4.2. Dependence on Inputs

The proposed algorithm is input-dependent, where the term "input" is used in a more general sense, referring to excitations, parameters, and initial conditions altogether. This subsection demonstrates that this explicit dependence can be both an advantage and disadvantage. It is a disadvantage in the sense that if the inputs are not chosen properly, the algorithm may yield a model that is simplified for a different scenario than the intended one. This is demonstrated on Example 1. Nevertheless, it is also advantageous to

have a tool that can take into account different scenarios and simplify the model accordingly. This is illustrated by Example 2.

Example 1: Importance of Proper Excitation

In this context excitations are considered proper if they correctly represent the scenario of interest, e.g., if they cause non-zero constraint forces for all constraints. Creating a proper excitation is a very important prerequisite step of the simplification algorithm, because an improper excitation may result in an overly-simplified model that is not suitable for the intended scenario, as the following example illustrates.

Consider the pendulum example given in Example 1 of Section 3.4.1 with the arbitrary choice of pendulum coordinate frame (Fig. 3.6a). The choice of the excitation becomes very critical in this case: if there is no excitation force in the model besides gravity, the simplification algorithm simplifies the original model (Fig. 3.7) as shown in Fig. 3.12. Note that there are three independent rotational DoF in the rigid-body model and no constraints, which means that the number of DoF of the simplified model is three, which is two more than the number of DoF of the original model.

The reason why the number of DoF changes after simplification is the improper excitation. Since there is no force that tries to move the pendulum away from its swing plane, the constraint forces get eliminated from the original model due to the inactivity of junctions, to which the constraint forces are connected. As long as the initial conditions satisfy the constraints there is no problem with the simplified model in terms of predicting the pendulum behavior. The problem is, however, that there is no mechanism left in the model to enforce the conformity of the initial conditions with the constraints.

The solution to this problem is to augment the original model with proper excitations. If we add to the original model forces that try to move the pendulum out of its swing plane, the simplified rotational joint model becomes as shown in Fig. 3.8.

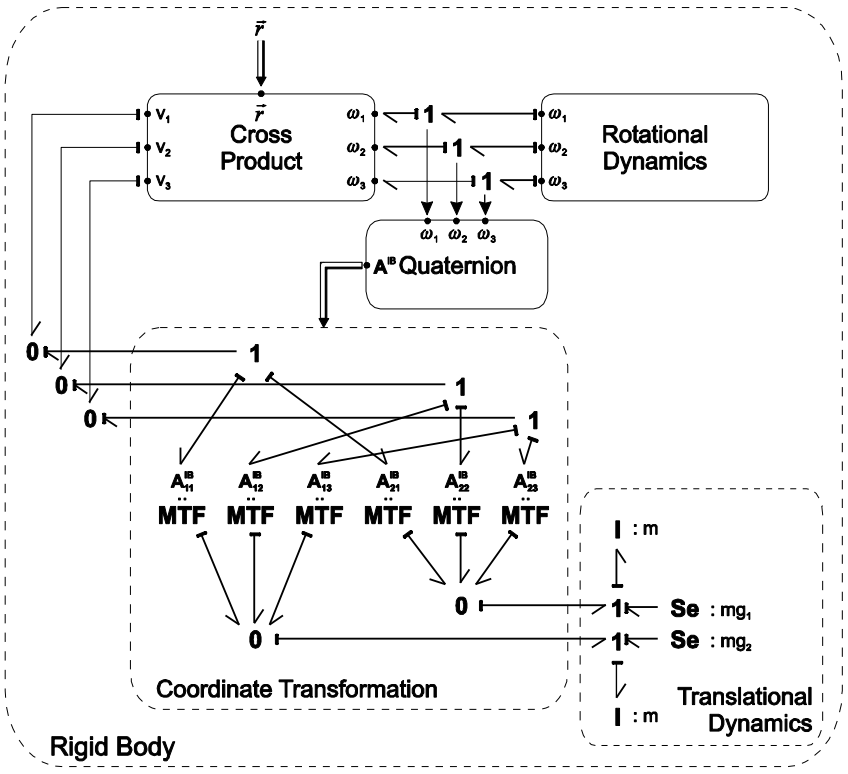


Figure 3.12. The simplified pendulum model obtained with a poor excitation

Although the rigid-body model still has three independent inertia elements, the system has only one DoF due to the two constraint elements kept in the rotational joint. Therefore, the model can not only prevent inconsistent initial conditions, but can also accommodate forces that do not lie in the swing plane.

A more general conclusion that can be drawn from this example is that a set of representative excitations instead of a single one may be necessary to describe the scenario of interest. In that case, the simplification has to be repeated for all excitations in the set. The simplified model for that particular scenario is then the union of the simplified models. In this work the selection of the set of excitations descriptive enough for a particular scenario is left to the modeler. Development of a tool to aid in this task is beyond the scope of this work, and could be an interesting future work.

Example 2: Advantage of Input-Dependence

The input-dependent characteristic of the proposed algorithm allows for the algorithm to take different scenarios into account. As a result, the algorithm can exploit the different scenarios and yield different simplified models. This example shows that the algorithm can simplify a model that is already simple enough for many different scenarios even further, when a particular input is considered.

Consider a system where two mass-spring systems are concatenated and the mass on the bottom is excited with a harmonic force, as given in Fig. 3.13, which also shows the bond graph of the system.

(a) (b)

Figure 3.13. (a) Two-mass-spring system, (b) its bond-graph model

The bond graph of the system may look simple enough, but for a certain parameter and initial condition set, the top mass-spring system can act like a perfectly-tuned passive vibration absorber. In that case, the bottom mass-spring system would not move, because the force generated by the vibration absorber would be symmetric to the applied force. If one applies the simplification procedure for this particular excitation, parameter and initial condition set, one obtains a simpler model, in which the inert mass-spring system is removed along with the excitation force (Fig. 3.14). In other words, two idle states and one excitation are removed from the model, and the system motion is then only due to the initial conditions of the remaining two states.

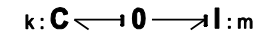


Figure 3.14. Simplified two-mass-spring system for the perfectly-tuned vibration absorber scenario

The general conclusion drawn for the previous example also applies here, i.e., the simplification may have to be repeated for different inputs to find a simplified model suitable for the scenario considered. For example, in the case of parameter uncertainty it may be a good practice to perturb the parameters to check if the simplification results are not for a very special case such as a perfectly tuned vibration absorber. Similarly, if a possible set/range of parameters exists, the simplification should be repeated for a representative set of parameters and the union of the results should be taken as the simplified model. How the representative parameters should be selected is an important issue that goes beyond the scope of this work, and addressing it is left as a future work.

3.5. SUMMARY AND CONCLUSION

Inactivity of a junction element is defined, and a bond-graph-level structural simplification procedure is proposed that is based on junction inactivity. This approach can be considered as a generalization of simplifying a bond graph by removing zero-flow 1-junctions and zero-effort 0-junctions. The suggested use of the activity metric allows for a unified treatment of junction elements and leads to a unified structural simplification procedure that is easy to implement at bond-graph level. Although the procedure is presented in a bond-graph framework, it can be easily implemented in other modeling environments (block diagrams, Modelica, etc.) as well. The procedure is illustrated with examples, and its realization-preserving and input-dependent properties are highlighted. Even though the examples in this chapter were selected from the mechanical domain, the proposed algorithm is applicable to any energetic domain.

The realization-preserving property of the proposed algorithm helps preserve the physical meaning of the model during simplification, but also raises the question about

which realizations to use to achieve best results. To some extent, this question is answered in the next chapter, and further investigation of this topic is left as a future work.

The input-dependent characteristic of the proposed algorithm, on the other hand, has been shown to be both a pitfall and an advantage. It is important to select the inputs (excitations, parameters, initial conditions) carefully, as they define the scenario that is being subject to simplification. Ill-defined scenarios may lead to overly simplified results, which is a potential pitfall. However, when the inputs are selected properly, the input-dependence of the algorithm can create different simplified models for different scenarios. From this point of view, this characteristic is considered as an advantage.

If more than one set of inputs is necessary to define the scenario of interest, e.g., in the case of stochastic inputs, parameter uncertainty, a range of possible parameter values, etc., then repetition of the simplification algorithm for all inputs is necessary. In that case, the union of the simplification results should be taken as the simplified model for that particular scenario. Proper selection of inputs is left to the modeler in this work. However, it may be a challenging task, and a tool to address this challenge could be valuable. This is another potential area for future work.

The results obtained so far with the proposed procedure encourage the integration of this tool with the modular modeling approach. Instead of creating a simple model by hand, which is error-prone and time-consuming, system models can be assembled quickly in a modular way, and then simplified automatically to increase insight and efficiency. This provides an alternative way to obtaining simple models, not just at the equation level, but also at the bond-graph level, while preserving all the benefits of modularity. This in turn can be very valuable from both a conceptual and computational point of view by allowing for rapid development of mathematical models tailored for specific scenarios of interest.

CHAPTER 4

ORIENTING BODY COORDINATE FRAMES USING KARHUNEN-LOÈVE EXPANSION FOR MORE EFFECTIVE STRUCTURAL SIMPLIFICATION

4.1. INTRODUCTION

Chapter 3 developed a junction-inactivity-based structural simplification technique that is particularly suitable for bond-graph models obtained through a modular approach. The technique is highly sensitive to the orientation of the body coordinate frames in multibody systems: improper alignment of body coordinate frames may prohibit a significant simplification. This chapter demonstrates how the Karhunen-Loève expansion can be used to automatically detect the existence of and to find the transformation into body coordinate frames that render the bond graph of a multibody system more conducive to simplification. The proposed technique is demonstrated using the simple example of a 3D pendulum constrained to move in a plane, but is applicable to arbitrarily complex multibody dynamics problems. The conclusion is that the Karhunen-Loève expansion successfully complements the junction-inactivity-based structural simplification technique when multibody dynamics are involved in the system and thus significantly contributes to the development of an automated modular modeling environment.

4.2. MOTIVATION

As discussed in the previous chapter, modular modeling, although beneficial especially for relatively easy creation, verification and management of models for complex systems, potentially lacks efficiency in two aspects: modular models may not be

as succinct as possible (conceptual inefficiency), which in turn may require more computational resources than necessary (computational inefficiency) [9]. The reason for these potential inefficiencies is the high level of detail included in the general-purpose submodels, which may become unnecessary for specific purposes in a particular system configuration. For example, the rotational dynamics in a 3D rigid-body submodel become redundant when the submodel is used to represent a point mass.

Eliminating such redundancies from a model to increase its conceptual and computational efficiency is highly desired. However, as systems become more complex, identifying the phenomena irrelevant for a specific scenario of interest and simplifying the system model becomes an important challenge, requiring much time and expertise. An automated model simplification algorithm would hence be an invaluable tool to address this challenge. Towards this end, previous chapter has shown that when the bond-graph language [21] is used to express the models, it is possible to simplify modular models at the bond-graph level systematically based on a concept called "junction inactivity" [2]. Such a systematic simplification complements the modular modeling approach well by increasing the conceptual and computational efficiency of the models obtained modularly [2].

One characteristic of the junction-inactivity-based simplification is that it does not change the realization of a given model. For mechanical systems, for example, this means that the original coordinates of a given model are not changed by the simplification algorithm. However, this makes the simplification sensitive to the realization of the model, because some realizations might be more conducive to simplification than others, as shown in previous chapter. Thus, the choice of realization is important for an effective simplification.

As a specific case of realization, this chapter focuses on the choice of body coordinate frames used in multibody systems. Certain orientations of a body coordinate frame, e.g., when axes are aligned with motions/constraints, can be more conducive to

simplification than others. For example, it is preferred to align the coordinate frame of a pendulum with the axis of rotation, so that only one rotational coordinate is necessary [2]. It is desired to find such preferred orientations systematically to benefit most from the simplification.

Such reorientation of coordinate frames changes the model's realization. However, because the change of realization is limited to coordinate frame reorientation and is thus performed only locally, the overall structure of the model continues to correspond to the structure of the physical system. Therefore, the model continues to be "physically meaningful" and thus intuitively appealing.

Reorienting coordinate frames is a special case of the problem of selecting a suitable set of independent coordinates to solve a set of differential algebraic equations (DAE), which has been studied widely in literature under the keywords "coordinate partitioning" and "tangent/null space methods". In particular, these keywords refer to the approach, in which, given a set of coordinates and constraints, a set of independent coordinates are sought to avoid integrating dependent coordinates and causing constraint violations. Kane's method, for example, provides a way to formulate the equations using an independent set of coordinates, and as a result the constraint forces with no virtual work are eliminated from the model [221, 222]. The selection of the independent set of coordinates, however, is left to the modeler.

Various approaches exist to automatically select the independent coordinates, including selecting a subset of the original coordinates using the LU decomposition [223], or creating independent linear combinations of the original coordinates by using the zero eigenvalues theorem [224], QR decomposition [225], singular value decomposition [226-228], or Gram-Schmidt orthogonalization [229-231].

These methods, including Kane's method, are based on the same fundamental concept that in a constrained mechanical system the motion evolves in the null space of the constraints, and differ only in the way they generate a basis for that lower

dimensional manifold [232]. After a basis is found, the equations are projected onto the lower dimensional manifold to obtain an independent set of coordinates.

The methods mentioned above have three limitations within the context of this chapter: (1) They depend on the existence of constraint equations. Hence, they are not readily applicable if models are formulated using realistic joints in a way such that there are no algebraic constraint equations, or if the motion evolves in a lower dimensional space not due to some constraints, but merely due to a particular selection of inputs and initial conditions. (2) With the exception of the LU decomposition [223], these methods create new coordinates by combining *all* of the original coordinates, thereby potentially losing their intuitive appeal. (3) These methods typically do not take structural simplification into account.

The search for a preferred coordinate frame can also be considered as a search for a subspace other than the full motion-space itself that captures the entire dynamics. If such a subspace exists, then the body coordinate frame should be aligned with the basis of that subspace to allow for simplification in the orthogonal complement of that subspace.

A well-established technique to identify dominant subspaces for a space of observations is the Karhunen-Loève expansion (KLE) [113, 114], which was reviewed in detail in Chapter 2. It has been widely used to *reduce* models by observing the dynamics of a system and projecting the equations of motion onto the dominant subspaces of the system dynamics. The main idea of this chapter, however, is to recognize that if instead of a *dominant* subspace, a subspace capturing the *entire* dynamics is sought, the KLE could solve the problem of orienting a given body coordinate frame to achieve a more significant *simplification*.

This chapter proposes a technique that utilizes the KLE to check if a preferred orientation of a coordinate frame exists, and, if it does, to obtain the transformation into the preferred coordinate frame. Thus, in this chapter the KLE is shown as a means to

achieve a more significant *simplification*, rather than *reduction*, by looking for preferred orientations of coordinate frames. In that sense, the usage of the KLE in this chapter is novel. Furthermore, the proposed technique furnishes a model more conducive to simplification regardless of the way the constraints are implemented (algebraic equations or parasitic elements), or even if the lower dimensional motion is due to a specific set of inputs and initial conditions rather than some constraints. In addition, because the change in realization is restricted to body coordinate frames, the proposed technique preserves the physical meaning of the model. Finally, the proposed method also considers the model structure, and can potentially yield orientations that yield a better simplification not only in dynamics, but also in kinematics. Hence, the proposed method alleviates the three aforementioned limitations of the coordinate partitioning methods in this context.

4.3. ORIENTING COORDINATE FRAMES USING THE KARHUNEN-LOÈVE EXPANSION

The general problem that this chapter is concerned with is to find better orientations for body coordinate frames to better set up a given model for simplification. The proposed procedure, however, is independent of the number of rigid bodies involved in the system, because for a multibody system the procedure is applied to each rigid-body independently. Hence, without loss of generality, the procedure is presented for a single rigid-body in this section.

As a building block of multibody systems, and as a generic representation of the models that are of interest to this chapter, consider the generic bond-graph representation of a 3D rigid-body with one connection point shown in Fig. 4.1 (for details please see Appendix A). In this representation the rotational dynamics are expressed in the body coordinate frame to achieve a constant inertia matrix, and the translational dynamics are expressed in the inertial frame.

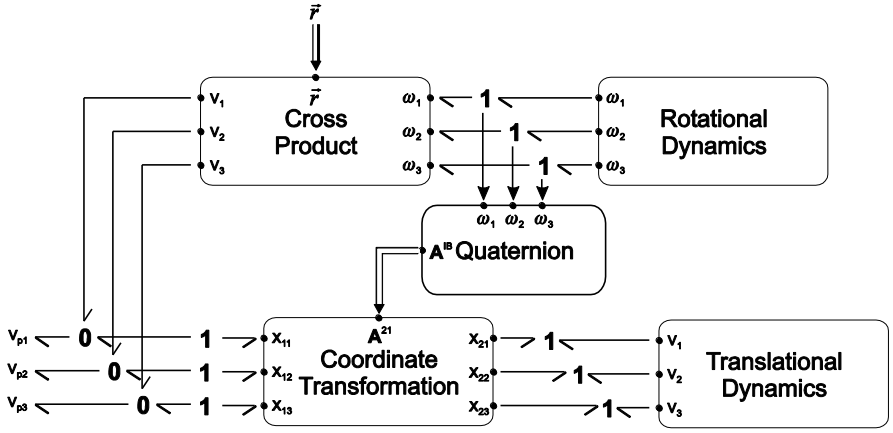


Figure 4.1. Bond-graph model of a 3D rigid body

Given the generic model in Fig. 4.1, it is desired to check if a certain orientation of the body coordinate frame is preferred for simplification, and, if so, to find the corresponding coordinate transformation into the preferred orientation. After the coordinate-frame reorientation, the junction-inactivity-based simplification can be applied to the model to obtain the simplified model. The flowchart in Fig. 4.2 outlines this generic procedure.

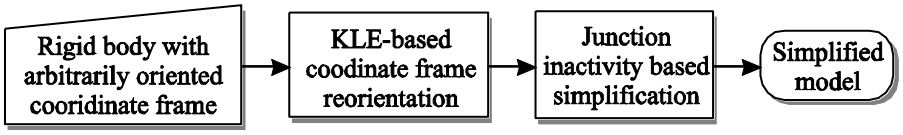


Figure 4.2. Flowchart of the simplification process

4.3.1. The Proposed KLE-Based Method for Orienting Coordinate Frames

Chapter 2 outlined the traditional way of using the KLE for model reduction purposes. This subsection presents a KLE-based method to reorient body coordinate frames for model simplification purposes. To this end, it is shown how the observation matrix **S** needs to be selected, and how the KLE results are interpreted and utilized.

First, let us determine the variables to be observed. A rigid body can do pure rotation, pure translation, or a combination of both. Thus, in general, 3 angular velocity

and 3 translational velocity components are necessary to define its motion. Therefore, to observe the full motion-space, the observation matrix **S** is formed as

$$\mathbf{S} = \left[\vec{v}_x \quad \vec{v}_y \quad \vec{v}_z \quad \vec{\omega}_x \quad \vec{\omega}_y \quad \vec{\omega}_z \right]_{m\times 6} = \left[\mathbf{S}_v \quad \mathbf{S}_\omega \right] \tag{4.1}$$

where $\vec{v}_x$, $\vec{v}_y$, $\vec{v}_z$ and $\vec{\omega}_x$, $\vec{\omega}_y$, $\vec{\omega}_z$ are the time histories of the *x*, *y*, *z* components of the velocity of the center of mass and angular velocity of the body, respectively, and the *xyz*-frame refers to the body coordinate frame.

Note that even though in general 6 variables are necessary to define a rigid body's translational and angular velocity, the physical space in which the body can move is only 3-dimensional. A KLE analysis with the full **S** matrix would result in 6-dimensional basis vectors, which would exceed the dimension of the physical space. Hence, the observation matrix is partitioned into translational and angular velocity observation, each with dimension $m\times 3$, as suggested by the right-hand side of Eq. (4.1). The KLE analysis is then performed separately, yielding 3-dimensional basis vectors in each case, i.e.,

$$\mathbf{S}_v = \mathbf{U}_v \mathbf{\Sigma}_v \mathbf{W}_v^T, \ \mathbf{S}_\omega = \mathbf{U}_\omega \mathbf{\Sigma}_\omega \mathbf{W}_\omega^T \tag{4.2}$$

Then, there exists a preferred orientation for the translational (angular) velocity, if the square of at least one and at most two of the singular values σ_{vi} ($\sigma_{\omega i}$) are zero within a numerical tolerance, i.e., $\sigma_{v1}^2 \geq \sigma_{v2}^2 \geq \sigma_{v3}^2 \approx 0$ with $\sigma_{v1}^2 > 0$ ($\sigma_{\omega 1}^2 \geq \sigma_{\omega 2}^2 \geq \sigma_{\omega 3}^2 \approx 0$ with $\sigma_{\omega 1}^2 > 0$). This is because the σ_{vi}^2 ($\sigma_{\omega i}^2$) specify how much signal energy is captured by the corresponding basis vectors $\vec{w}_{vi}$ ($\vec{w}_{\omega i}$), and a zero σ_{vi}^2 ($\sigma_{\omega i}^2$) implies that there is no translation along (rotation about) the corresponding basis vector $\vec{w}_{vi}$ ($\vec{w}_{\omega i}$). Note, however, that if all σ_{vi}^2 ($\sigma_{\omega i}^2$) are zero, then the body is not translating (rotating) and thus the orientation of the body coordinate frame is immaterial for translation (rotation). Therefore, the singular values are indicators of the existence of a preferred frame.

Furthermore, the coordinate transformation matrix from the original body coordinate frame *B* into the preferred body coordinate frame P_v (P_ω) is given by the transpose of the matrix $\mathbf{W}_v$ ($\mathbf{W}_\omega$), i.e.,

$$\mathbf{A}^{P_v B} = \mathbf{W}_v^T, \ \mathbf{A}^{P_\omega B} = \mathbf{W}_\omega^T \tag{4.3}$$

Note that the preferred frame P_v will, in general, be different than P_ω. This means that the preferred coordinate frame to express translation will be, in general, different than the one to express rotation. One possible way of reconciling the two results is proposed in the next subsection.

Regardless of the way it is chosen, once a preferred coordinate frame is found, any vector $\vec{r}^{\,B}$ and the inertia matrix $\mathbf{J}^B$ expressed in the original body coordinate frame, and the initial value of $\mathbf{A}^{IB}$, $\mathbf{A}_0^{IB}$, can be transformed into the preferred coordinate frame using

$$\begin{aligned} \vec{r}^{\,P} &= \mathbf{A}^{PB}\vec{r}^{\,B} \\ \mathbf{J}^P &= \mathbf{A}^{PB}\mathbf{J}^B\left(\mathbf{A}^{PB}\right)^T \\ \mathbf{A}_0^{IP} &= \mathbf{A}_0^{IB}\left(\mathbf{A}^{PB}\right)^T \end{aligned} \tag{4.4}$$

respectively.

4.3.2. The Proposed Reconciliation Algorithm

An algorithm is presented in this subsection to reconcile the different results one may obtain from the rotational and translational KLE analyses. It is worth noting at the outset that the proposed algorithm is just one way of reconciling the results, and may not be optimal. It is given here just to show one possible way of reconciliation, and creating an algorithm that will yield the best possible orientation for maximum simplification is beyond the scope of this work.

The proposed algorithm is based on several observations and assumptions. These can be summarized as follows: First, note that both the angular velocity, $\vec{\omega}$, and the rotational dynamics are expressed in the body coordinate frame. In bond graph terms, the I-elements representing the rotational dynamics are directly connected to the 1-junctions representing the components of $\vec{\omega}$. This implies that the removal of a component of $\vec{\omega}$ will result in the removal of the rotational dynamics in the corresponding direction. Therefore, eliminating a component of $\vec{\omega}$ directly implies simplification in rotational

dynamics. In other words, if the rotational KLE analysis is indicating a preferred orientation, it is guaranteed that the preferred frame will give simplification in rotational dynamics.

On the other hand, note that the translational velocity, *v*, is expressed in the body coordinate frame, but the translational dynamics are expressed in the inertial coordinate frame. In bond graph terms, the I-elements representing the translational dynamics are not directly connected to the 1-junctions representing the components of *v*. This means that it is not possible to achieve any more simplification in translational dynamics by reorienting the body coordinate frame. Thus, any simplification in dynamics achieved through reorientation is going to be in the rotational domain. In other words, the translational KLE analysis can yield simplification only in kinematics and, if at all, in rotational dynamics.

It is also important to note that the generic rigid body model considered in this work accepts only orthonormal bases. Oblique bases are beyond scope. Finally, let us assume that simplification in dynamics is preferred over simplification in just kinematics.

Under these observations and assumptions, the following algorithm is proposed to reconcile the results obtained from the rotational and translational analyses:

1. Perform KLE on both $\vec{\omega}$ and $\vec{v}$. Let the rotational and translational singular values be $\sigma_{\omega i}$ and σ_{vi}, $i = 1,2,3$, respectively. Let the corresponding basis vectors be $\vec{w}_{\omega i}$ and $\vec{w}_{vi}$, $i = 1,2,3$, respectively.

2. Because of observations mentioned above, it is first desired to align the body frame to eliminate as many $\vec{\omega}$-components as possible. Hence, $\sigma_{\omega i}$ are considered first. Four cases are possible:

   a. $\sigma_{\omega i} > 0, i = 1,2,3$. This implies that the body is doing a fully 3D rotation and there is no preferred frame as far as the rotation is concerned. In this

case the coordinate frame should be reoriented according to the translational analysis, if it is suggesting a preferred orientation.

b. $\sigma_{\omega 1} \geq \sigma_{\omega 2} > \sigma_{\omega 3} = 0$. This implies that two axes are enough to describe the body's rotation, namely, $\vec{w}_{\omega 1}$ and $\vec{w}_{\omega 2}$, and these are the first two basis vectors for the preferred coordinate frame. The third axis is given by the cross product $\vec{w}_{\omega 1} \times \vec{w}_{\omega 2}$, and thus the preferred frame is uniquely defined.

c. $\sigma_{\omega 1} > \sigma_{\omega 2} = \sigma_{\omega 3} = 0$. This implies that the body rotates about one axis only, namely, $\vec{w}_{\omega 1}$, and this is the first basis vector of the preferred coordinate frame. The preferred orientation for ω is not yet uniquely defined: any vector orthogonal to $\vec{w}_{\omega 1}$ could be used as the second preferred basis vector to uniquely define the preferred orientation as far as the simplification in the rotational domain is concerned. Thus, the translational analysis is considered to see if it is possible to take advantage of this remaining degree of freedom in the reorientation:

   i. If it exists, the $\vec{w}_v$ with zero σ_v that is orthogonal to $\vec{w}_{\omega 1}$ is taken as the second preferred basis vector. The third preferred basis vector is then defined by the cross product of the first two. If there are two $\vec{w}_v$ that satisfy the above condition, the choice is arbitrary, because the cross product will automatically give the second orthogonal $\vec{w}_v$. As a result, both $\vec{w}_v$ are included in the preferred set of basis vectors.

   ii. If such a $\vec{w}_v$ does not exist, use the $\vec{w}_{\omega 2}$ and $\vec{w}_{\omega 3}$ to fully determine the preferred coordinate frame.

d. $\sigma_{\omega i} = 0, i = 1,2,3$. This implies that there is no rotation. In this case the coordinate frame should be reoriented according to the translational KLE analysis, if it is suggesting a preferred orientation.

Because of the rigid-body representation considered in this work, this algorithm gives preference to the rotational KLE analysis over the translational one. It gives an orientation that will yield maximum simplification rotational dynamics, and if the preferred frame is not uniquely defined by the rotational KLE analysis, it attempts to take advantage of the translational KLE analysis to achieve more simplification in kinematics. The benefit of reorienting the body coordinate frame to simplification is highlighted in the next section on an example.

4.4. EXAMPLE: PENDULUM

As an illustrative example, consider a simple pendulum with an arbitrarily oriented body coordinate frame (Fig. 4.3), which is modeled modularly using 3D rigid-body and joint models (Fig. 4.4).

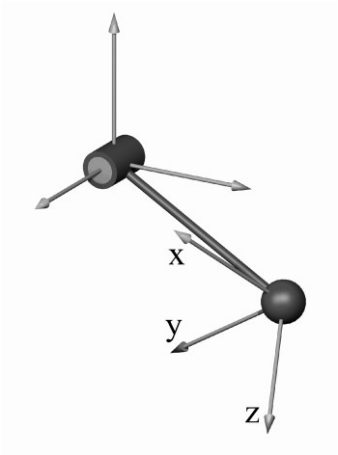


Figure 4.3. Pendulum with an arbitrarily oriented body coordinate frame

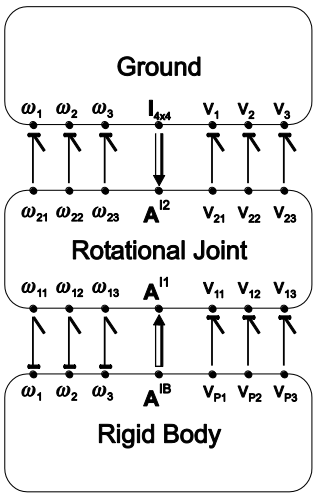


Figure 4.4. The modular bond-graph model of the pendulum

The parameter values and initial conditions for the model are summarized in Table 4.1. With the given orientation of the pendulum coordinate frame, the simplification procedure presented in the previous chapter yields the model in Fig. 4.5 for an inactivity threshold of 10^{-5}, which is set equal to the integration tolerance. Note that due to the arbitrary orientation of the body coordinate frame all rotational inertia elements are kept in the model along with the two constraints represented by the PSf elements. The simplification in the translational dynamics is a result of the fact that the translational dynamics are expressed in the inertial frame. When resolved in the inertial

Table 4.1. Parameter values and initial conditions for the original pendulum model

Parameters	**Values**
Gravity	$\vec{g}^{\,I} = \begin{bmatrix} 0 & -9.81 & 0 \end{bmatrix}^T \text{ m/s}^2$
Pendulum mass	$m = 1 \text{ kg}$
Pendulum inertia	$\mathbf{J}^B = 0.004\, \mathbf{I}_{3\times 3} \text{ kg} \cdot \text{m}^2$
Connection point	$\vec{r}^{\,B} = \begin{bmatrix} 0.66724 & 0.48260 & -0.56735 \end{bmatrix}^T \text{ m}$
Rotation axis	$\vec{a}^{\,B} = \begin{bmatrix} 0.66904 & -0.05353 & 0.74130 \end{bmatrix}^T \text{ m}$
Initial conditions	**Values**
Angular momentum	$\vec{H}_0^B = \begin{bmatrix} 8.028 & -0.642 & 8.896 \end{bmatrix}^T 10^{-3} \text{kg} \cdot \text{m}^2/\text{s}$
Position	$\vec{p}_0^{\,I} = \begin{bmatrix} 0 & -1 & 0 \end{bmatrix}^T \text{ m}$
Euler parameters	$\left(e_0, \vec{e}\right)_0 = \left(0.79864, 0.16084 \begin{bmatrix} 1 & -2 & 3 \end{bmatrix}^T\right)$

frame, which is aligned with the axis of rotation, the translational velocity has only two non-zero components, and the simplification reflects this. Note, however, when resolved in the body coordinate frame, the translational velocity has three non-zero components. Thus, the arbitrary orientation of the body coordinate frame prohibits further simplification.

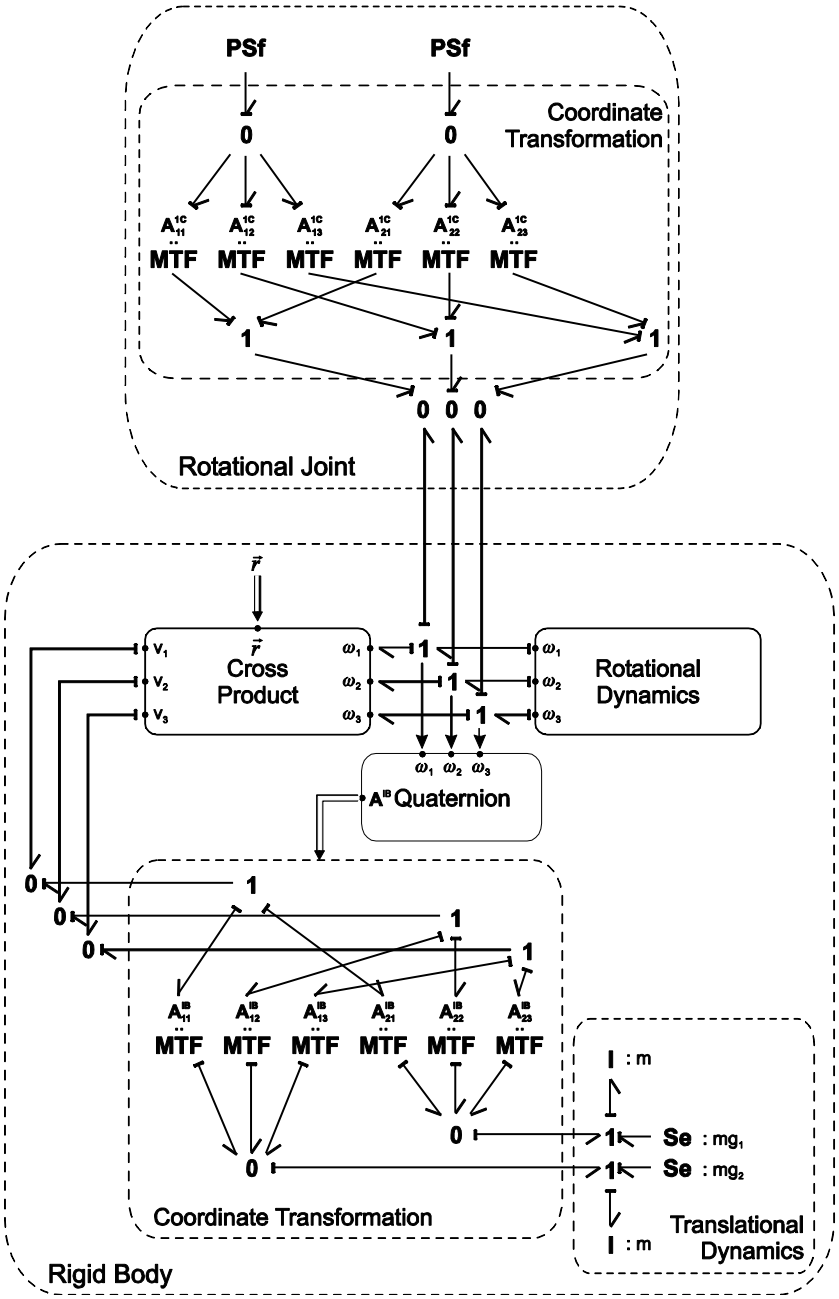


Figure 4.5. The simplified pendulum model for the arbitrary pendulum coordinate frame

To look for a better orientation of the pendulum coordinate frame the proposed KLE-based technique is applied. First, KLE is applied to both the translational and angular velocity observations, which give

$$\begin{aligned}\boldsymbol{\Sigma}_v^2 &= diag(1488.8, 0, 0);\ \mathbf{W}_v = \begin{bmatrix} -0.3274 & -0.6690 & -0.6672 \\ 0.8742 & 0.0535 & -0.4826 \\ 0.3586 & -0.7413 & 0.5673 \end{bmatrix} \\ \boldsymbol{\Sigma}_\omega^2 &= diag(1488.8, 0, 0);\ \mathbf{W}_\omega = \begin{bmatrix} -0.6690 & 0.2410 & -0.7031 \\ 0.0535 & -0.9279 & -0.3690 \\ -0.7413 & -0.2845 & 0.6079 \end{bmatrix}\end{aligned} \tag{4.5}$$

The existence of two zero $\sigma_{\omega i}^2$ indicates that there is a preferred body coordinate frame where the angular velocity expressed in the body coordinate frame will have only one non-zero component. Using the proposed reconciliation algorithm one obtains the set $\{\vec{w}_{\omega 1}, \vec{w}_{v3}, \vec{w}_{\omega 1} \times \vec{w}_{v3} = \vec{w}_{v1}\}$ as basis vectors of the preferred coordinate frame, i.e., the transformation matrix into the preferred frame is given by

$$\mathbf{A}^{PB} = \begin{bmatrix} -0.6690 & -0.6672 & -0.3274 \\ 0.0535 & -0.4826 & 0.8742 \\ -0.7413 & 0.5673 & 0.3586 \end{bmatrix}^{\mathrm{T}} \tag{4.6}$$

The preferred frame is illustrated in Fig. 4.6.

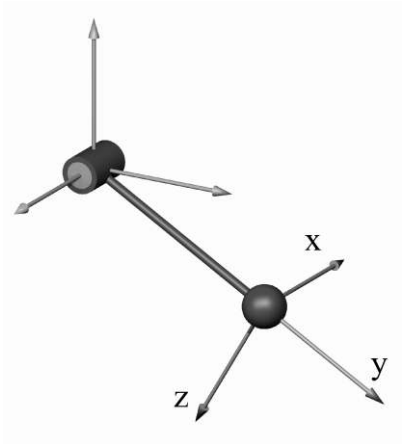


Figure 4.6. The preferred pendulum coordinate frame

Reorienting the pendulum coordinate frame as suggested by $\mathbf{A}^{PB}$, and applying the junction-inactivity-based algorithm, much more dramatic simplification results are achieved, as shown in Fig. 4.7. This result is due to the fact that in the preferred frame one of the axes (x) is aligned with the axis of rotation, thereby reducing the number of necessary rotational inertia elements to one, and eliminating the need for constraints. Furthermore, another axis (y) is aligned with the pendulum. Therefore, in this frame, the translational velocity has only one non-zero component, which leads to a better simplification. The new values for the parameters and initial conditions that are affected by the reorientation are given in Table 4.2.

Table 4.2. Parameter values and initial conditions affected by the reorientation

Parameters	**Values**
Connection point	$\vec{r}^{\,B} = \begin{bmatrix} 0 & -1 & 0 \end{bmatrix}^T$ m
Rotation axis	$\vec{a}^{\,B} = \begin{bmatrix} -1 & 0 & 0 \end{bmatrix}^T$ m
Initial conditions	**Values**
Angular momentum	$\vec{H}_0^B = \begin{bmatrix} -0.012 & 0 & 0 \end{bmatrix}^T$ kg·m²/s
Euler parameters	$(e_0, \vec{e})_0 = \left(0, 0.707107 \begin{bmatrix} -1 & 0 & 1 \end{bmatrix}^T\right)$

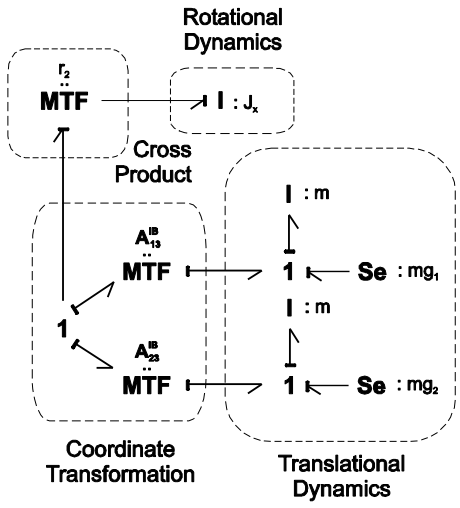


Figure 4.7. The simplified pendulum model for the reoriented pendulum coordinate frame

4.5. DISCUSSION

4.5.1. Structural Simplification

The example given in Section 4.4 shows that by using the proposed KLE-based technique it is possible to detect the existence of, and find the transformation into a coordinate frame that is more conducive to simplification. The fact that aligning one of the axes of the pendulum frame with the axis of rotation yields a better simplification is revealed by the KLE analysis. The KLE analysis further reveals that aligning another axis with the pendulum, some additional simplification in kinematics can be achieved. Specifically, if one of the axes is aligned with the axis of rotation, but the orientation is arbitrary otherwise, as is the case with the frame given by the rotational KLE analysis, the simplified model given in Fig. 4.8 is obtained. Notice that while the simplification in dynamics is the same as in Fig. 4.7, the simplification in kinematics is not to the same extent. Thus, while looking for a better orientation the proposed method not only

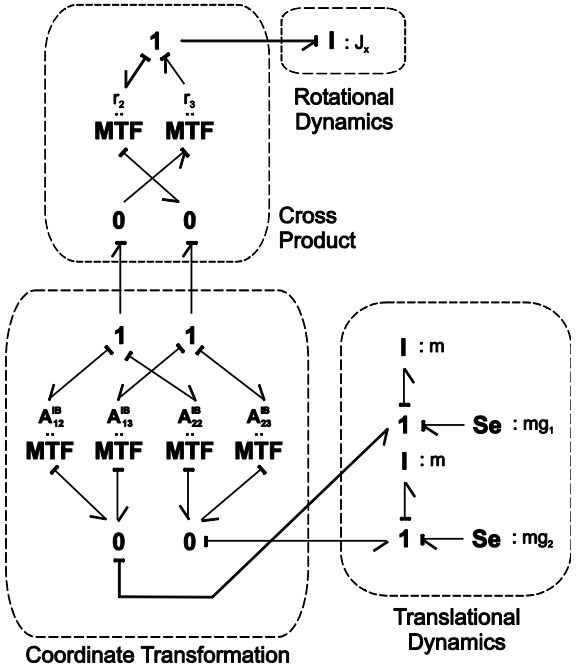


Figure 4.8. The simplified pendulum model for the P_ω frame

considers the dynamics, but also the kinematics. Since the procedure is systematic, it is possible to automate it and reduce the need for human insight.

The pendulum example illustrates a case in which the translational and rotational KLE analyses give different preferred coordinates. The proposed reconciliation algorithm is successfully utilized to synergistically combine the results. Note that in this example the resulting preferred coordinate frame is equivalent to P_v. However, this is a special case, and in general the resulting frame does not necessarily have to be equivalent to P_v or P_ω.

4.5.2. Local Change in Realization

Note that even though the proposed technique results in a change in realization of the model, the physical meaning of the model is preserved. This is because the proposed technique applies KLE only locally, limiting the changes in realization to reorientation of body coordinate frames. Furthermore, recall that the simplification algorithm applied after the reorientation is realization-preserving. Thus, the simplified model obtained by applying the proposed coordinate-frame-reorientation technique followed by the junction-inactivity-based simplification is still physically meaningful and intuitively appealing.

4.5.3. Realistic Joints

Since the proposed technique relies on the observation of the motion and not on constraint equations, it can be readily applied to multibody systems with realistic joints, which can be formulated without any constraint equations. Consider the pendulum example given in Section 4.4, but assume that the ideal rotational joint block is replaced with a more realistic implementation, as shown in Fig. 4.9. This implementation considers the stiffness and damping of the joint, as well as the friction in the axis of rotation. In addition to the original parameter values considered in Table 4.1, let the joint

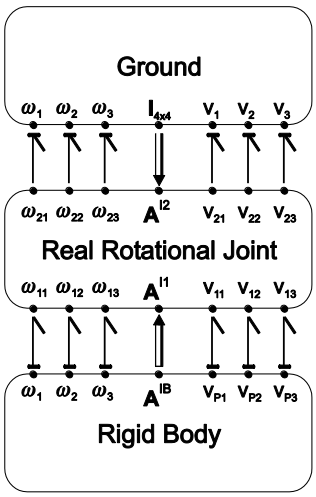


Figure 4.9. The modular bond-graph model of the pendulum with a realistic joint

parameters be $k_{\omega i} = 100\ \text{N}\cdot\text{m/rad}$, $b_{\omega 1} = b_{\omega 2} = 100\ \text{N}\cdot\text{m}\cdot\text{s/rad}$, $b_{\omega 3} = 0.1\ \text{N}\cdot\text{m}\cdot\text{s/rad}$, $k_{vi} = 500\ \text{N/m}$, and $b_{vi} = 500\ \text{N}\cdot\text{s/m}$ (please see Appendix A for details).

In this case the proposed KLE analysis yields the following preferred frame

$$\mathbf{A}^{PB} = \begin{bmatrix} -0.6690 & -0.6678 & -0.3261 \\ 0.0535 & -0.4810 & 0.8751 \\ -0.7413 & 0.5680 & 0.3575 \end{bmatrix}^{\mathrm{T}} \tag{4.7}$$

which is very close to the frame obtained with the ideal joint (Eq. (4.6)). When this preferred frame is used, the simplification algorithm yields the model in Fig. 4.9. Note that the realistic joint becomes partially idealized by the simplification algorithm, as the PSf element in the joint block indicates. Also note that the extent of simplification is less in this case due to the considered realistic joint effects. Nevertheless, this example demonstrates that the proposed method can be successfully applied to realistic joints as well.

4.5.4. Further Remarks

It is important to note that after the proposed application of the KLE it is known which of the 1-junctions representing the translational or angular velocity components, if any, are going to be inactive. Thus, those 1-junctions could be removed without the aid of

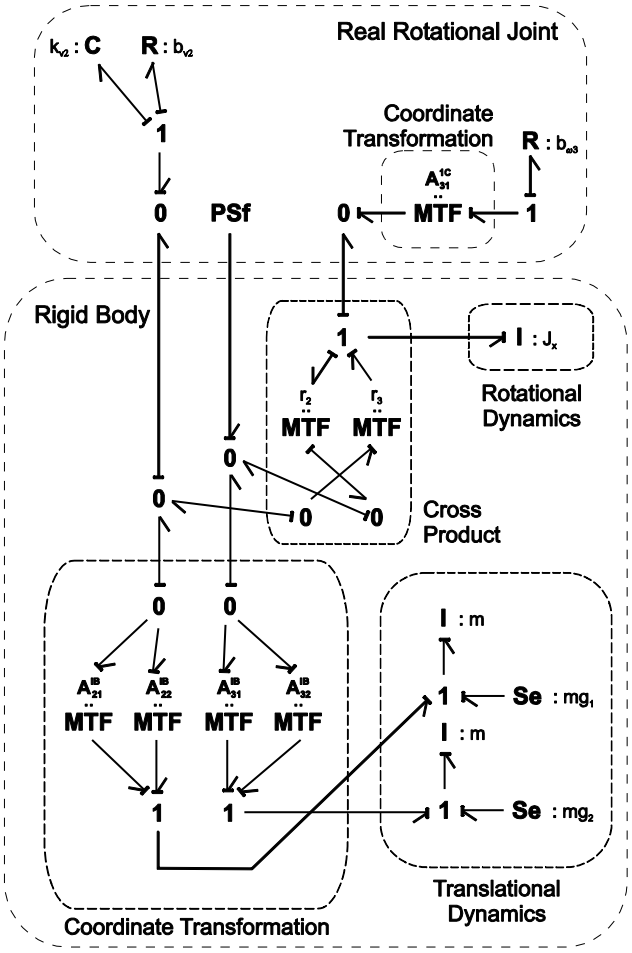


Figure 4.10. Pendulum model with a realistic joint simplified for the reoriented pendulum frame

the simplification algorithm. However, this would allow for only a limited amount of simplification, because the usage of the KLE in the context of this chapter is very local. The application of the simplification algorithm presented in [2] considers the model globally including the non-mechanical domains, thus yielding better results.

It is also worth noting that once the simplified model in Fig. 4.7 is obtained, it can be further simplified by using some of the well-known bond-graph simplification rules. Specifically, the model in Fig. 4.11 can be obtained from the model in Fig. 4.7 by lumping the dependent inertias onto the independent one, eliminating the zero-effort source mg_1, and lumping the cross product and the coordinate transformation into the effort source mg_2 to create a modulated effort source. It would be very difficult,

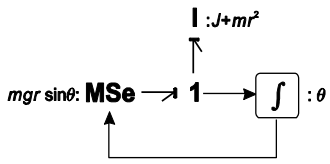


Figure 4.11. Further simplified pendulum model

however, to obtain the model in Fig. 4.11 starting from the original model in Fig. 4.4 and using only these simple rules. Therefore, the proposed KLE-based technique followed by the junction-inactivity-based simplification algorithm complements successfully the well-known bond-graph simplification rules.

Finally, the pendulum example demonstrates the fundamental steps behind the envisioned automated modular modeling environment, in which a system is first modeled modularly, and then automatically simplified to tailor the model for specific scenarios of interest. As the example shows, the proposed KLE-based technique increases the efficiency of the simplification algorithm, and is therefore an important contribution to the development of the automated modular modeling environment.

4.6. SUMMARY AND CONCLUSION

A Karhunen-Loève-expansion-based technique is proposed to detect the existence of and find the transformation into coordinate frames that are more conducive to simplification in systems involving multibody systems. The proposed method allows the Karhunen-Loève expansion to be used in a novel way as a means for bond-graph level simplification, rather than its traditional usage for equation-level reduction. The idea is demonstrated on a simple pendulum example, in which the initial arbitrary orientation of the pendulum coordinate frame is systematically reoriented to achieve a more significant simplification.

The conclusion is that when multibody systems are part of a complete system, the proposed technique complements the simplification algorithm given in the previous

chapter well by reorienting coordinate frames automatically and increasing the effectiveness of the simplification. This, in turn, reduces the need for human insight, and saves time in creating and simulating complex systems involving multibody systems. Therefore, the proposed method also contributes significantly to the development of an automated modular modeling environment.

Future work includes creating reconciliation methods for rigid-body representations other than the one considered in this work, such as one that allows for oblique frames, or expresses translational dynamics in the body coordinate frame as well. Future work may also explore the possibility to automatically look for alternative realizations for a given model beyond just reorienting coordinate frames and find the best physically-meaningful realization that yields the most significant simplification.

CHAPTER 5

REALIZATION-PRESERVING STRUCTURE AND ORDER REDUCTION OF NONLINEAR ENERGETIC SYSTEM MODELS USING ENERGY TRAJECTORY CORRELATIONS

5.1. INTRODUCTION

Previous chapters developed algorithms for *simplifying* the structure of a lumped dynamic system model. This chapter moves further to enable simultaneous model *structure* and *order reduction*. Specifically, it introduces a new energy-based metric to evaluate the relative importance of energetic connections in a model. This metric (1) accounts for *correlations* between energy flow patterns in a model using the Karhunen-Loève expansion; (2) examines *all* energetic connections in a model, thereby assessing the relative importance of both energetic *components* and their *interactions*, and enabling both order and structural reduction; and (3) is realization-preserving, in the sense of not requiring a state transformation. A reduction scheme based on this metric is presented and illustrated using a simple example.

5.2. MOTIVATION

As discussed in Chapter 1, in modeling dynamic systems, one often faces a tradeoff between model accuracy on the one hand and model simplicity and tractability on the other. Models that balance these conflicting requirements for their respective applications are deemed *proper* by the literature [12].

Obtaining such proper models can be challenging, because it is difficult to know at the outset which phenomena are important to model and which can be safely neglected.

One solution to this problem involves first modeling as many phenomena as possible, with an implicit preference to fidelity over simplicity, then *reducing* the resulting models till they are proper.

As reviewed in detail in Chapter 2, the literature presents many *model reduction* techniques including component cost analysis [170], balanced truncation [141], polynomial approximation methods [90, 94, 96], optimal Hankel norm approximation [105], optimization-based methods [174, 175], proper orthogonal decomposition [117, 128], singular perturbation [57], and modal analysis [67, 68].

The conclusion of the review in Chapter 2 is, however, that each of these techniques has one or more of the following limitations: (1) Applicability to a limited set of systems (e.g. linear, time-invariant systems, systems with multiple time-scales, etc.); (2) requiring a realization change; (3) trajectory independence; (4) being limited to equation level; (5) not considering the structure of the model.

A model reduction algorithm is sought in this chapter that addresses the above limitations. Specifically, an algorithm is sought that is applicable to nonlinear systems, realization-preserving, trajectory-dependent, applicable at the graph level, and also aimed at both structure and order reduction. In developing this algorithm, the focus is on *energetic* systems, i.e., those composed of components that store, dissipate, and exchange energy. Such systems are quite common in engineering. Furthermore, one can potentially construct reduction algorithms for such systems that utilize the domain-independent, intuitively appealing notions of *energy* and *power* as foundations for model reduction.

Chapter 2 shows that the literature presents several examples of *energy-based* model reduction algorithms. For instance, Rosenberg and Zhou utilize root-mean-square power flow in an energetic interconnection to assess its importance [197]. This furnishes a trajectory-dependent and realization-preserving model structure and order reduction algorithm that can be applied directly to a given model's graph representation.

Similarly, Louca *et al.* propose to use the normalized $\mathcal{L}_1$-norm of power flow through an energetic component (e.g., mass, spring, damper, etc.) as a metric for model reduction [20]. They call this metric the *activity index* of the given component, and further propose a Model Order Reduction Algorithm (MORA) based on it [20]. This method is also trajectory-dependent, realization-preserving, and applicable directly to a given model's graph representation. However, as its name suggests, it concentrates on reducing model order, rather than model structure.

Rideout *et al.* extend the activity metric to systematically detect *decoupling* among the elements of a model and *partition* it accordingly [19]. The resulting algorithm is trajectory-dependent, realization-preserving, and applicable to graph representations of nonlinear systems, but it focuses only on model partitioning, and does not consider a reduction beyond partitioning.

Finally, previous chapters present a technique for detecting energetically inactive junction structures in a given model and *simplifying* the model accordingly. The proposed algorithm simplifies both the order and structure of the given model, and can be applied to graph-level nonlinear models without requiring a realization change. It does not, however, enable reduction beyond the simplification threshold.

This chapter extends the above energy-based reduction literature by developing a new reduction metric that takes into account not only the magnitudes of the various energy trajectories in a system, but also their correlations. Intuitively, the algorithm seeks to determine not only which components and interconnections are most active, but also which ones affect overall system behavior the most. Towards this goal, the algorithm applies the Karhunen-Loève expansion, a correlation analysis technique, to the various trajectories of energy flow through the various interconnections in a dynamic system. It then projects the results of the Karhunen-Loève expansion back onto the original system realization to quantify the "importance" of the various components and interconnections in this realization. This furnishes an energy-based, realization-preserving, and trajectory-

dependent model structure and order reduction algorithm that is applicable to the graph representations of nonlinear systems.

5.3. PROPOSED METRIC AND REDUCTION METHOD

Any reduction technique is based on a metric to evaluate which phenomena can be neglected in a given model. Hence, this section first introduces a new metric that accounts for the correlations between the energy flow patterns in a model. The motivation behind creating this new metric is twofold: First, because of the fact that the energy flow patterns in a system determine the system's behavior, it is hypothesized that energy is a natural choice for the basis of a new metric. Such an energy-based metric would allow for a unified treatment of not only different energy domains (e.g., mechanical, electrical, hydraulic, etc.), but also the dynamic components and their interactions in a system. Second, it is further hypothesized that taking into account the correlations between the energy flow patterns throughout the model would allow for a better assessment of the relative importance of each part of the model to the whole response. Based on these two hypotheses it is proposed that the Karhunen-Loève expansion be combined with the energy-based approach in a new way to assess the energy exchange phenomena in a model relatively. After the new metric is introduced, a realization-preserving reduction algorithm is proposed based on this metric.

For ease of presentation, the bond-graph language [21] will be used in the rest of this chapter. Since the bonds in a bond graph represent energetic connections, the bond-graph representation is convenient for the calculations presented below. Furthermore, although the metric is not limited to bond graphs, having a metric that is directly applicable to bond graphs is advantageous, because the graphical representation is a higher-level representation than mathematical equations, which means that bond-graph

level simplifications are automatically reflected in the mathematical equations. The opposite, however, is not necessarily true.

With that in mind, the proposed metric is developed as follows. In the first step, the Karhunen-Loève expansion (KLE) is applied to the energy trajectories of the bonds in a bond-graph model with the goal of capturing the correlations between them. In particular, let the energy trajectory of each bond in a bond-graph model be arranged column-wise in a matrix **S**, i.e.,

$$\mathbf{S} = \left[\vec{E}_1 \quad \vec{E}_2 \quad \dots \quad \vec{E}_n \right]_{m \times n} \tag{5.1}$$

where n is the number of bonds in the bond graph, m is the number of observations, and $\vec{E}_i$ is the energy trajectory of the i^{th} bond. Singular value decomposition of **S** yields:

$$\mathbf{S} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \tag{5.2}$$

where $\mathbf{\Sigma} = diag(\sigma_1, \sigma_2, \dots, \sigma_n)_{m \times n}$ with $\sigma_1 \geq \dots \geq \sigma_n \geq 0$. It is established that the columns of the orthogonal $n \times n$ **V** matrix form a basis for the observation space, and the squares of the singular values provide a measure of how much signal energy is captured by each basis vector [128]. Since the observed quantity is energy, the columns of **V** give a new basis to express the energy flow in the system, so they can be interpreted as energy exchange modes in the system. Within each mode, the (absolute value of the) j^{th} component tells how much j^{th} bond contributes to that mode. Furthermore, the squared singular values give a measure of how much each mode contributes to the observed response. Thus, in the second step, a measure for the importance of the bonds is obtained by a weighted combination of the absolute values of the modes, where the weights are the squared singular values, i.e.,

$$\vec{I} \triangleq \sum_{i=1}^{n} \sigma_i^2 \left| \vec{v}_i \right| \tag{5.3}$$

where $\vec{v}_i$ is the i^{th} column of **V**, and $\vec{I}$ is the importance vector of the bonds, whose j^{th} component gives the importance of bond j. The importance vector can be normalized with respect to its maximum element to give a relative measure of importance, i.e.,

$$\overrightarrow{RI} = \vec{I} / \max(\vec{I}) \tag{5.4}$$

Based on this proposed metric, a model reduction algorithm can be outlined as follows:

1. Simulate the model, and record the energies of the bonds of interest.

2. Arrange the data in a matrix such that the columns are the energy trajectories of the corresponding bonds, i.e., $\hat{\mathbf{S}} = \left[\vec{E}_1 \quad \vec{E}_2 \quad \dots \quad \vec{E}_n \right]_{m \times n}$.

3. Calculate $\mathbf{S}_{i,j} = \sqrt{\Delta t_i}\, \hat{\mathbf{S}}_{i,j}\,; i = 2,\dots,m;\ j = 1,\dots,n$, where Δt_i is the time step between the i-1st and the i^{th} observations. This step is to account for unevenly spaced observations.

4. Perform singular value decomposition on $\mathbf{S}$, i.e., $\mathbf{S} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$.

5. Calculate the relative importance of the bonds using Eq. (5.3) and (5.4).

6. Arrange bonds in decreasing order of relative importance *RI*. Let p be an index for the rows of this ordered list.

7. If $RI_p / RI_{p+1} > r$ for some row $p < n-1$ and user-defined ratio $r > 1$, then bonds in rows $p+1,\dots,n$ are subject to reduction. There may be more than one such threshold, i.e., more than one level of reduction. It is up to the modeler to decide on the ratio r and which threshold to use for reduction.

8. Remove the elements that got disconnected from the rest of the model as a result of step 7.

Note that if all bonds are subject to the analysis, this algorithm gives a unified approach to the reduction problem in the sense that not only the order, but also the structure of the model can be reduced. This will be hereafter referred to as a global

application of the metric. It is also possible to perform the analysis locally, e.g., only for the bonds connected to the components representing the states for the purposes of model order reduction, or only for the bonds connected to a junction element for the purposes of model partitioning. Both the global and local applications of the proposed method are demonstrated in the next section on an example.

5.4. ILLUSTRATIVE EXAMPLE

This section provides an example to illustrate the mechanics of the proposed method and emphasize its advantages. The example system is first reduced globally for two different scenarios. This shows the proposed method's applicability to nonlinear systems, ability to achieve graph-level reduction, and ability to reduce the order and structure of the model, while taking into account the scenario of interest and preserving the realization of the model. Then, the analysis is performed locally for the second scenario to compare the proposed method to MORA, thereby stressing the benefit of having a metric that considers the correlations between the energy flow patterns in a system.

Consider the system shown in Fig. 5.1, where a mass-spring-damper system is connected to the slider of a crank mechanism. A rotational spring and damper are connected to the crank arm, and the rotational spring is undeflected when $\alpha = \pi/2$. There is viscous friction between the slider and ground. The parameter values are given in Table 5.1 and the bond graph model of the system is given in Fig. 5.2, which will be

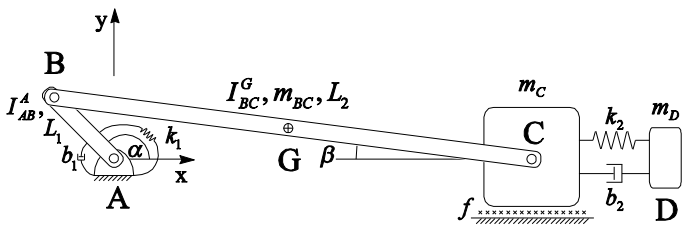


Figure 5.1. Schematic representation of the example system

Table 5.1. The parameters of the example system

Parameter	Value	Parameter	Value
I_{AB}^{A}	3.53×10^{-5} kg·m^2	k_2	100 N/m
I_{BC}^{G}	8.84×10^{-3} kg·m^2	f	1 N·s/m
m_{BC}	0.42 kg	b_1	0.01 N·m·s/rad
m_C	10 kg	b_2	0.1 N·s/m
m_D	0.1 kg	L_1	0.05 m
k_1	1 N·m/rad	L_2	0.5 m

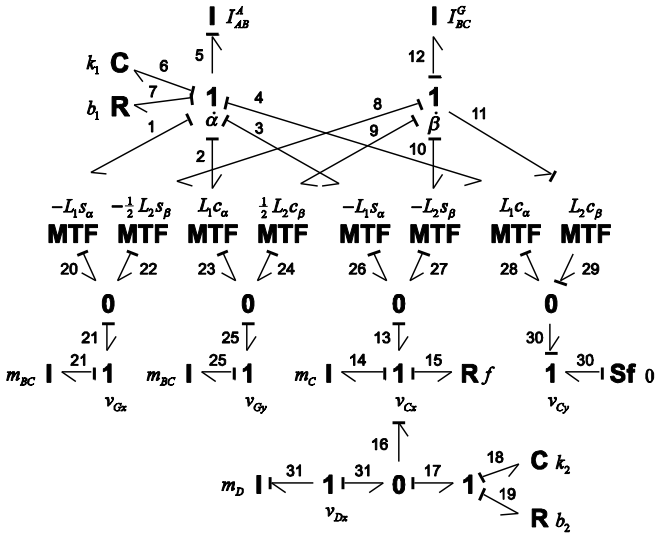


Figure 5.2. Bond graph of the example system

hereafter referred to as the full model. The bonds are numbered such that each bond with a unique energy receives a unique index. Bonds connected to power-through junctions therefore have the same index. The full model includes the dynamics of the links and masses, as well as the kinematics

$$\begin{aligned}
\vec{v}_G &= \vec{v}_B + \vec{\omega}_{BC}\times\vec{r}_{G/B} = \vec{\omega}_{AB}\times\vec{r}_{B/A} + \vec{\omega}_{BC}\times\vec{r}_{G/B} \\
&= \left(-L_1\dot{\alpha}\sin\alpha - \frac{L_2}{2}\dot{\beta}\sin\beta\right)\vec{i} + \left(L_1\dot{\alpha}\cos\alpha + \frac{L_2}{2}\dot{\beta}\cos\beta\right)\vec{j} \\
&= v_{Gx}\vec{i} + v_{Gy}\vec{j} \\
\vec{v}_C &= \vec{v}_B + \vec{\omega}_{BC}\times\vec{r}_{C/B} = \vec{\omega}_{AB}\times\vec{r}_{B/A} + \vec{\omega}_{BC}\times\vec{r}_{C/B} \\
&= \left(-L_1\dot{\alpha}\sin\alpha - L_2\dot{\beta}\sin\beta\right)\vec{i} + \left(L_1\dot{\alpha}\cos\alpha + L_2\dot{\beta}\cos\beta\right)\vec{j} \\
&= v_{Cx}\vec{i} + v_{Cy}\vec{j}
\end{aligned} \tag{5.5}$$

along with the constraint

$$v_{Cy} = L_1\dot{\alpha}\cos\alpha + L_2\dot{\beta}\cos\beta = 0 \tag{5.6}$$

5.4.1. Scenario 1 – Global Analysis

Consider the scenario in which the springs k_1 and k_2 are given initial displacements of 1 rad and 0.01 m, respectively, where a positive sign indicates extension, and the free response of the system is observed. Let the output of interest be the position of the mass m_D. When the proposed method is applied to this scenario, the results summarized in Table 5.2 are obtained for a global relative importance analysis with $r = 2$ and a simulation time-window of 5 seconds.

The dashed lines in Table 5.2 indicate the 5 thresholds for $r = 2$, and hence 5 different levels of reduction, which can be explained physically as follows:

Level 1: This threshold points to a well-known structural simplification that can be made in the bond graph; namely, the null flow-source can be removed along with the 1-junction v_{Cy} without affecting the accuracy of the model.

Level 2: The moment of inertia of the second link is removed. Even though it is

Table 5.2. Relative importance of bonds for Scenario 1 and reduction thresholds for $r = 2$

Bond	Relative importance		Bond	Relative importance	
7	100%		1	1.86%	
26	64.93%		21	1.69%	
3	64.93%		17	1.37%	
13	56.60%		31	0.91%	
14	39.84%		18	0.48%	3
6	27.11%		22	0.21%	
15	21.25%	5	8	0.21%	
11	10.25%		23	0.13%	
29	10.25%		2	0.13%	
28	10.25%		5	0.08%	
4	10.25%		25	0.07%	
27	10.02%		24	0.07%	
10	10.02%	4	9	0.07%	2
19	2.79%		12	0.02%	1
16	1.90%		30	0	
20	1.86%				

larger than the moment of inertia of the first link, the second link goes through a much less amount of rotation, and therefore the energy associated with its rotational dynamics is very low.

Level 3: The rotational and translational dynamics of the first link, and the translational dynamics and kinematics of the second link in y-direction are removed. Furthermore, the translational kinematics of the point G in x-direction is reduced by neglecting the terms involving β and its derivatives, i.e. the expression for $\vec{v}_G$ in Eq. (5.5) reduces to

$$\vec{v}_G = -L_1\dot{\alpha}\sin\alpha\vec{i} \tag{5.7}$$

Level 4: The dynamics and kinematics of the mass-spring-damper system connected to the slider are removed, as well as the translational dynamics and kinematics of the second link in x-direction.

Level 5: The kinematics associated with β are removed. As a result, the expression for $\vec{v}_C$ in Eq. (5.5) reduces to

$$\vec{v}_C = -L_1\dot{\alpha}\sin\alpha\vec{i} \tag{5.8}$$

and the constraint Eq. (5.6) is not needed. Figure 5.3 shows the schematic representation of this reduced system, and Fig. 5.4 shows the corresponding bond graph. Figure 5.5 compares the output of this reduced model to the output of the full model.

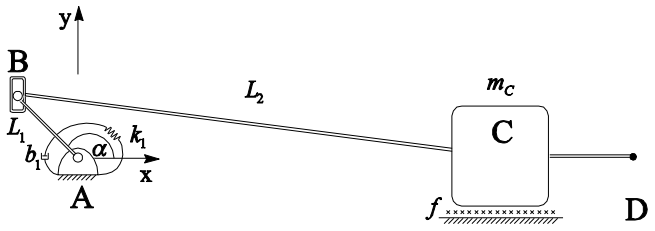


Figure 5.3. Schematic representation of the 5th-level reduced model for Scenario 1

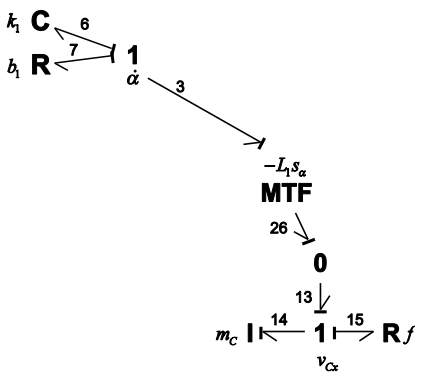


Figure 5.4. Bond graph of the 5th-level reduced model for Scenario 1

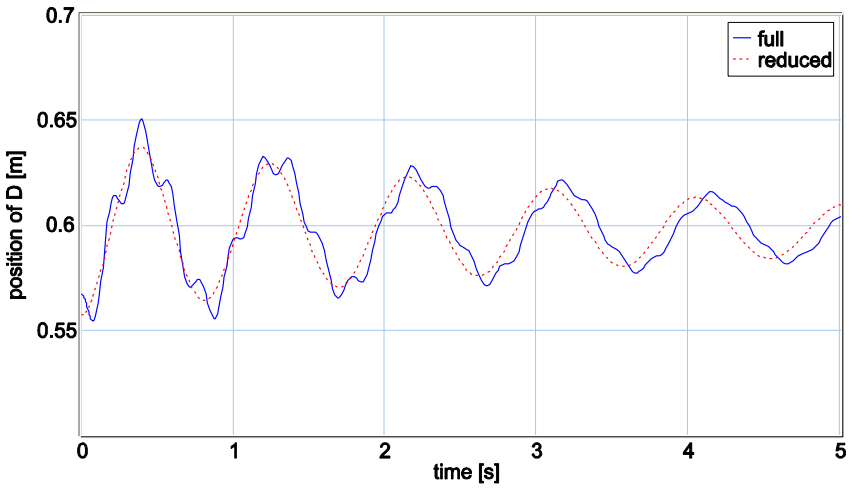


Figure 5.5. Output of the full model vs. the 5th-level reduced model for Scenario 1

So far this example illustrated the mechanics of the global application of the proposed metric and highlighted the following benefits of it: applicability to nonlinear systems, ability to achieve graph-level reduction, preservation of the original realization of the model, and ability to reduce the structure of the model, i.e. reducing not only the dynamics, but also the kinematics. The next part highlights the metric's ability to furnish different reduced models for different scenarios.

5.4.2. Scenario 2 – Global Analysis

Consider now the scenario in which an initial momentum of $0.05\ \text{kg}\cdot\text{m/s}$ is given to the mass m_D, and the free response of the system is observed. As in Scenario 1, let the output of interest be the position of the mass m_D. When the proposed method is applied to this scenario, the results summarized in Table 5.3 are obtained for a global relative importance analysis with $r = 2$ and a simulation time-window of 3 seconds.

Table 5.3. Relative importance of bonds for Scenario 2 and reduction thresholds for $r = 2$

Bond	Relative importance		Bond	Relative importance	
17	100%		21	0.02%	1
19	85.70%	4	4	0.00%	
18	21.93%		29	0.00%	
31	19.22%	3	28	0.00%	
16	1.81%		11	0.00%	
13	1.14%		10	0.00%	
3	1.14%		27	0.00%	
26	1.14%		22	0.00%	
7	0.83%		8	0.00%	
14	0.48%		2	0.00%	
6	0.35%		23	0.00%	
15	0.21%	2	25	0.00%	
5	0.03%		24	0.00%	
20	0.02%		9	0.00%	
1	0.02%		12	0.00%	
			30	0	

Table 5.3 indicates 4 thresholds and hence 4 different levels of reduction, however since the level 4 corresponds to practically discarding the model completely, it will be ignored. The remaining reduction levels can be explained physically as follows.

Level 1: All kinematics involving β, the translational kinematics and dynamics of the second link in y-direction as well as its rotational dynamics, and the constraint Eq. (5.6) are removed. Without any kinematics involving β, the expression for $\vec{v}_G$ in Eq. (5.5) reduces to Eq. (5.7).

Level 2: The rotational and translational dynamics of the first link, and the kinematics and dynamics of the second link in x-direction are removed.

Level 3: The kinematics involving α, the rotational stiffness k_1 and damping b_1, the kinematics and dynamics of mass m_D along with the friction between m_D and the surface are removed. Schematically, the system reduces down to Fig. 5.6, and the bond graph reduces to Fig. 5.7. Figure 5.8 compares the output of this reduced model to the output of the full model.

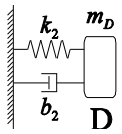


Figure 5.6. Schematic representation of the 3rd-level reduced model for Scenario 2

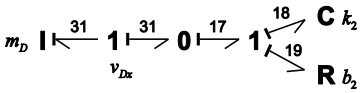


Figure 5.7. Bond graph of the 3rd-level reduced model for Scenario 2

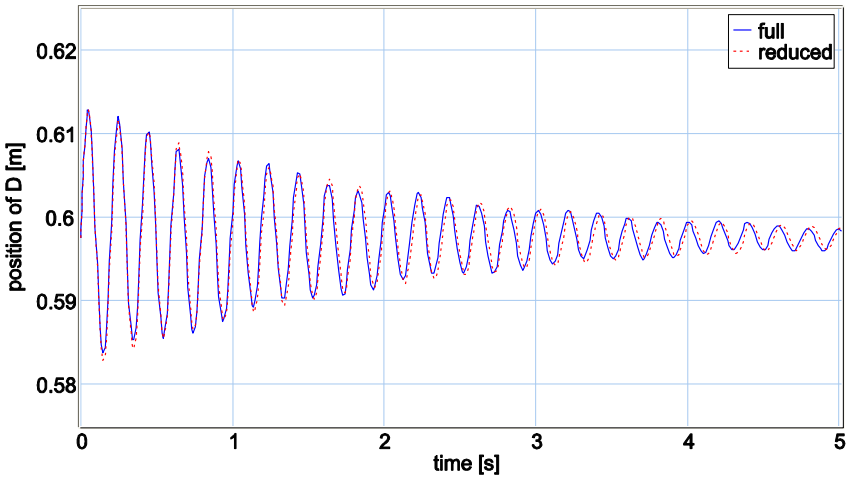


Figure 5.8. Output of the full model vs. the 3rd-level reduced model for Scenario 2

Notice the difference between the reduced models for the two scenarios (Fig. 5.4 and 5.7), and how the proposed method tailors the reduction according to the scenario of interest.

5.4.3. Scenario 2 – Local Analysis

The last part of this example applies the proposed metric only to bonds connected to I, C and R elements to look for possibilities of reduction in dynamics only. This serves two purposes: first, the local application of the proposed metric is illustrated; and second, this makes it possible to compare the proposed method to MORA, and emphasize the benefit of taking into account the correlations between the energy flow patterns.

When the I, C and R elements in the full model are assessed with both relative importance and activity, the results summarized in Table 5.4 are obtained. Notice the difference between the rankings of the elements, especially in the top three rows. The reason for this difference is twofold: First, MORA works with activity, a weighted $\mathcal{L}_1$-norm of power, whereas the proposed method works directly with energy. That means that the activity of an element always increases, even though its energy might decrease, and that may cause activity to overestimate the significance of some elements. Second, MORA uses only the final value of activity to rank the components, while the proposed method makes use of the entire history of energy flow.

As a result, notice how low the activity index of the R element b_2 (bond 19) is, to the extent that it looks like b_2 can be removed from the model without affecting the response too much. If MORA is used to preserve, for example, 96% of the total activity in the system (threshold 2 in Table 5.4), b_2 would be eliminated, and the response characteristic of the system would change dramatically. If, however, a slightly more conservative threshold, such as 97% (threshold 1 in Table 5.4), is used, MORA would keep b_2 in the reduced model. This high sensitivity to the threshold is a result of the low activity index of b_2 that is actually a quite important element.

On the other hand, the proposed metric ranks b_2 as the most important element, thereby keeping it in the reduced model regardless of the chosen threshold, and preserving the damped nature of the response. Therefore, this shows that taking into

Table 5.4. Comparing the proposed method to MORA

Proposed method			MORA		
Bond	Rel. Importance		Bond	Activity Index	
19	100%	4	18	48.35%	
31	27.36%		31	48.00%	2
18	27.13%	3	19	2.42%	1
7	0.98%		14	0.91%	
14	0.60%		6	0.24%	
6	0.40%		21	0.04%	
15	0.24%	2	7	0.03%	
5	0.03%		15	0.01%	
21	0.03%	1	5	0.00%	
25	0.00%		25	0.00%	
12	0.00%		12	0.00%	

account the correlations between the energy flow patterns in the system improves the assessment of what is negligible and what is not.

5.5. DISCUSSION

The proposed metric is not the first application of the KLE for model reduction purposes. In fact, KLE is the basis for some well-established reduction methods such as the proper orthogonal decomposition or balanced truncation. However, the proposed metric differs from the existing KLE-based methods in two ways: (1) It observes energy trajectories rather than state trajectories, which broadens the scope from state variables to any energetic connection, enabling the inclusion of system structure in reduction considerations. (2) As opposed to projecting the system onto the new basis given by $\mathbf{V}$ and using the singular values to differentiate between the important and unimportant states in this *new* realization, which may have no physical meaning, the proposed metric combines the modes and singular values to reflect back on the original realization and assess the relative importances of the bonds in the *given* realization.

As illustrated in the previous section, the proposed reduction algorithm has the following characteristics:

1. It not only reduces the order of the model, but also its structure. This unified treatment gives the proposed scheme a novel reduction perspective compared to the existing methods.

2. It preserves the realization of the original model. Therefore, if the original realization has a physical meaning of particular interest, that physical meaning will be preserved in the reduced model.

3. It is applicable to nonlinear models at graph-level.

4. It considers local and global applications of the proposed metric. The local application, where only some of the bonds in the model are included in the analysis, can be used for, e.g., model order reduction or model partitioning. The global application, which considers all bonds in the model, allows for concurrent reduction of the order and structure of the model.

5. It accounts for the input trajectories and initial conditions of interest, and furnishes different reduced models for different inputs and initial conditions.

Because of the trajectory-dependent nature of the proposed method, it is important to select the inputs and initial conditions carefully, such that they truly capture the scenario of interest. This might involve performing the importance analysis several times, if necessary, with different sets of inputs and/or initial conditions, and combining the results. Although the algorithm will yield a proper model for the considered set of inputs and initial conditions, it is up to the user to make sure that this set properly captures the scenario of interest.

It is also important to select carefully the initial values in the energy calculations. If a bond is connected to an energy storage element, the initial energy of that element should be taken into account in the energy calculation of the bond connected to it. In the

first scenario, for example, the springs have the initial potential energies $\frac{1}{2}k_1\left(\alpha_0-\frac{\pi}{2}\right)^2=0.5\text{ J}$ and $\frac{1}{2}k_2x_0^2=0.005\text{ J}$, and those initial energies are taken as the initial conditions for the energies in bonds 6 and 18, respectively.

In addition to initial conditions and input dependency, output dependency may also be a desired characteristic in a reduction algorithm, i.e., the algorithm should be able to tailor the reduction according to the outputs of interest. However, the proposed metric is not, in its current form, output dependent. This means that it may eliminate some low-energy parts of a system from the model, even if they are providing some outputs of interest. Although it is possible to avoid that by keeping the model parts that are of particular interest out of the scope of analysis by applying the proposed method locally to the rest of the model, it is still desired to formally incorporate the different output weights of the different parts of the system into the method. This is left as future work.

5.6. SUMMARY AND CONCLUSION

A new energy-based metric is proposed to assess the importance of energetic interconnections in a nonlinear energetic system model. The proposed metric combines the energy-based and correlation-analysis-based approaches in a unique way, such that not only it improves the assessment of what is important and what is negligible *in the original realization*, but also makes it possible to include the model *structure* in the assessment. In terms of the bond-graph representation this corresponds to ranking *all* the bonds in the order of relative importance.

Based on this metric, a reduction algorithm is proposed that is applicable to nonlinear models at graph level, preserves the original realization, and simultaneously reduces the model *order* and *structure* based on the inputs and initial conditions of interest. An example is provided to illustrate the mechanics and highlight the benefits of the proposed method.

CHAPTER 6

CASE STUDY: PROPER MODELING OF THE HMMWV

6.1. INTRODUCTION

So far the algorithms developed have been demonstrated on simple textbook-type examples. This chapter aims to present a case study on a system that better reflects the complexity of contemporary engineering systems to illustrate the performance of the proposed algorithms on such systems. Specifically, the proper modeling of the Army's High Mobility Multipurpose Wheeled Vehicle (HMMWV) is presented in detail.

First, the multibody dynamics of the HMMWV is modeled through a modular approach using the 3D rigid body and joint models library previously developed by the author. This model is referred to as the "full" model.

Three different scenarios are then considered, for which proper models are sought: a two double-lane-change maneuver, shaker table, and driving straight. In the first scenario the vehicle is accelerated from rest, two double lane change maneuvers are performed, and the vehicle is brought back to rest. This scenario might be of interest for rollover studies. The second scenario represents a virtual shaker table testbed to study, e.g., the ride quality of the vehicle. In the third scenario the vehicle is accelerated, driven at constant speed, and decelerated to stop without any steering input. This scenario could be useful for studying the acceleration properties of the vehicle.

The full model is then simplified considering all three scenarios simultaneously, i.e., a single simplified model is obtained that retains the accuracy of the full model for all three scenarios. This is done by performing the junction-inactivity analysis for all

three scenarios and combining the results. The simplified model is then reduced for the three scenarios separately, obtaining a different reduced model for each scenario. As a result, different proper models of the same system are obtained that are tailored for different scenarios of interest and for different levels of fidelity.

6.2. MODULAR MODELING OF THE HMMWV

Previous work by the author has created the framework for a modular approach to modeling 3D multibody systems using the bond graph representation [9, 10]. Using this approach, as well as the information about the kinematic structure and some parameter values of the HMMWV available in literature [233-235], a multibody model of the HMMWV is created in this section.

The model includes: the chassis; the four suspensions comprising lower and upper A-arms, wheel hubs, suspension springs and dampers; the front and rear anti-roll bars; the four tires with tire stiffness and damping, and longitudinal and lateral slip models; and the steering mechanism consisting of the steering link, idler arm, Pitman arm, and tie rods. Figure 6.1 shows the components of the HMMWV considered in this case study. The anti-roll bars are not shown in the figure.

Modular modeling of these components is presented next in detail, and the HMMWV model is assembled from these component models. The parameter values are given in the Appendix B.

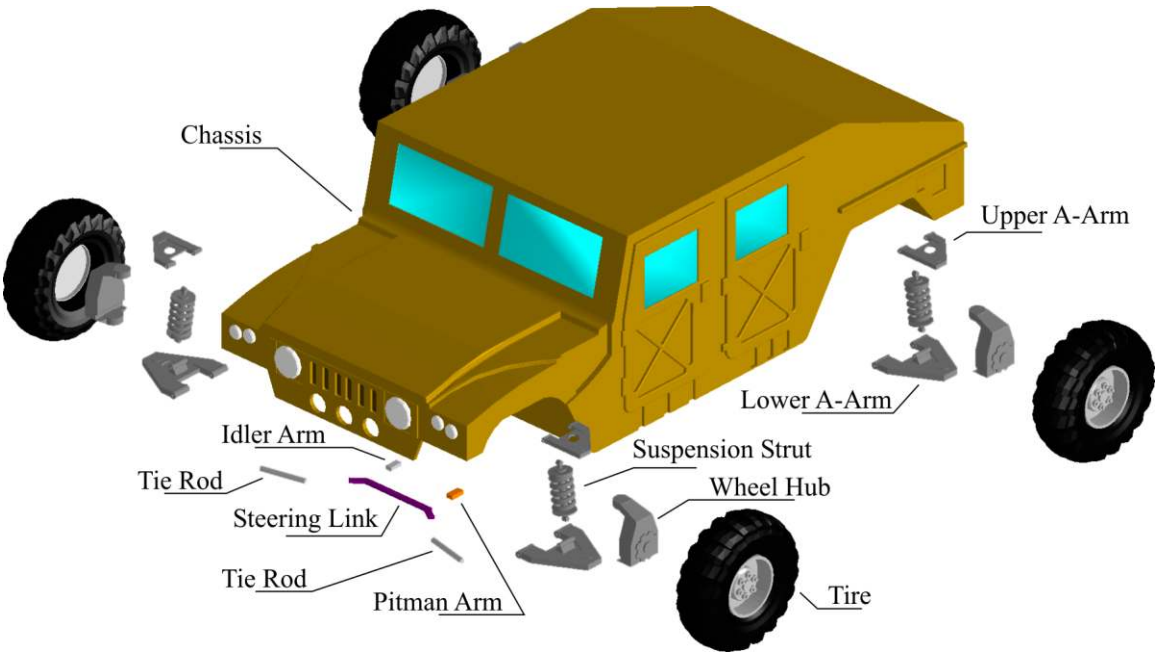


Figure 6.1. The components of the HMMWV

6.2.1. The Chassis

The chassis is modeled as a rigid body with connection points for the lower and upper A-arms, and the idler and Pitman arms. Figure 6.2 shows the chassis module. In the subscripts, FLU refers to the front left upper A-arm, RRL refers to the rear right lower A-arm, etc., and subscripts I and P stand for the Idler and Pitman Arms, respectively. For details of the rigid body model, please refer to the Appendix A.

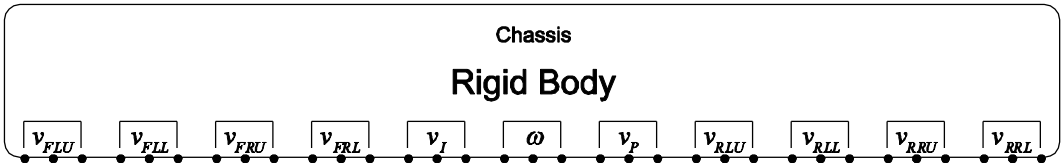


Figure 6.2. Model of the chassis

6.2.2. The Suspensions

The HMMWV has an independent suspension system both in front and rear, i.e., each wheel on the same axle can move independently of each other. The suspensions are composed of lower and upper A-arms, wheel hub, and suspension spring and damper. The A-arms and the wheel hub are modeled as rigid bodies that are connected through joint models as shown in the kinematic structure in Fig. 6.3. Note that the kinematic structure of the front and rear suspensions are different. Specifically, the wheel hubs are connected to the A-arms through rotational joints in the back, and through spherical joints in front to allow for steering. For the same reason, the front wheel hubs have an additional connection point where the tie rods of the steering mechanism are connected, but this interconnection is not shown in the figure.

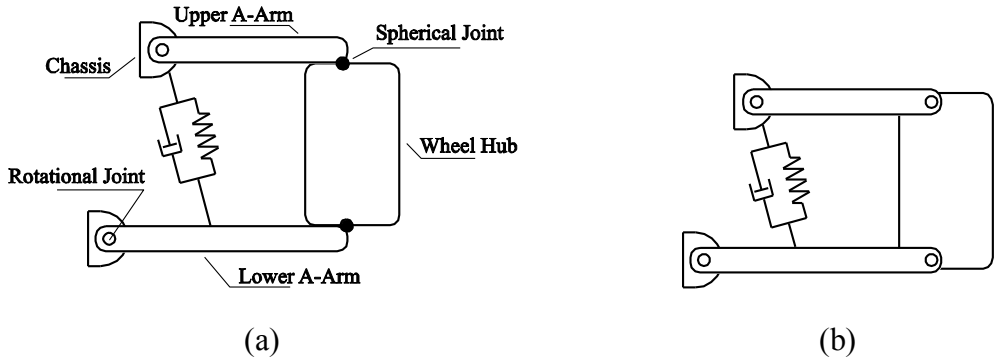


Figure 6.3. The suspension mechanism: (a) front; (b) rear

The suspension models for the front and back are given in Fig. 6.4 and 6.5, respectively. Note that when modeling the suspension springs and dampers, an auxiliary coordinate frame is introduced, whose z-axis is aligned with the suspension spring and damper.

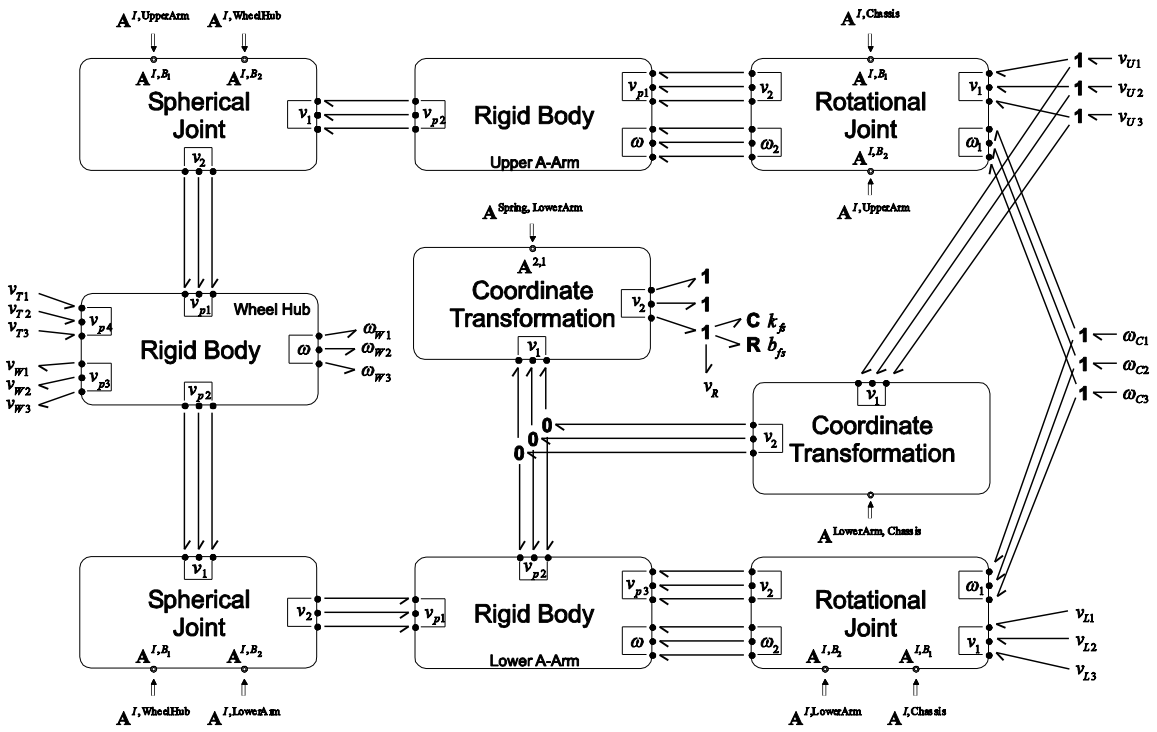


Figure 6.4. The front suspension model

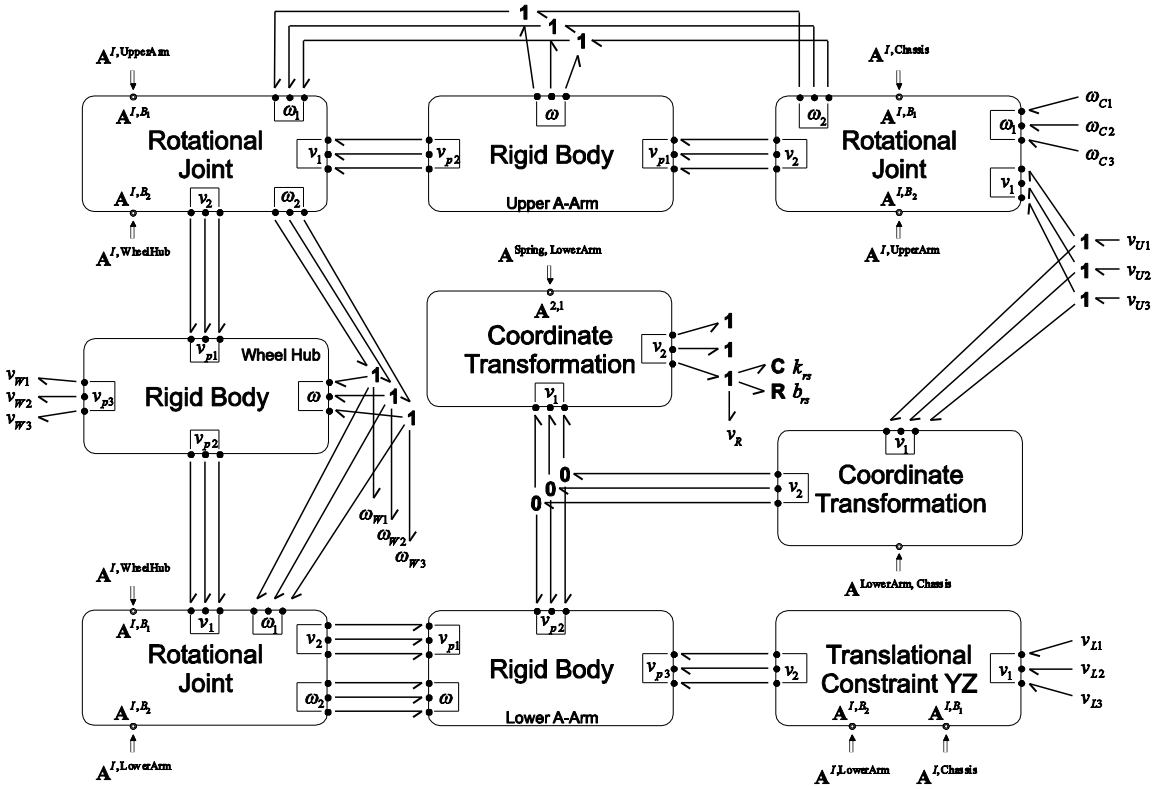


Figure 6.5. The rear suspension model

6.2.3. The Tires

The tires are modeled mainly as rigid bodies except that the tire stiffness and damping are taken into account. To this end, a coordinate frame C is introduced whose origin and y-axis coincide with those of the body-fixed frame, but the z-axis always points towards the contact point with the road, as illustrated in Fig. 6.6. The tire stiffness and damping act along the z-axis of this coordinate frame C. The multibody model of the tire is given in Fig. 6.7.

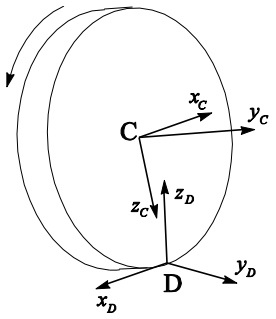


Figure 6.6. The auxiliary tire coordinate frames

The contact of the tire with the road is modeled through longitudinal and lateral slip models and a constraint equation. To aid with representing the tire-road interaction, another auxiliary coordinate frame, D, is introduced, also shown in Fig. 6.6. The origin of D is the contact point of the tire with the road. The z-axis is the normal of the road surface at the contact point, and the y-axis of D is the projection of the y-axis of the body-fixed frame onto the plane tangent to the road surface at the contact point. Thus, the x- and y-axes of D respectively define the longitudinal and lateral directions, which are necessary to express the slip models, and the z-axis defines the direction for the constraint force that keeps the contact point on the road surface. Figure 6.8 gives the tire and road model, in which the longitudinal and lateral slip models are implemented through R-elements, and the constraint through the PSf-element.

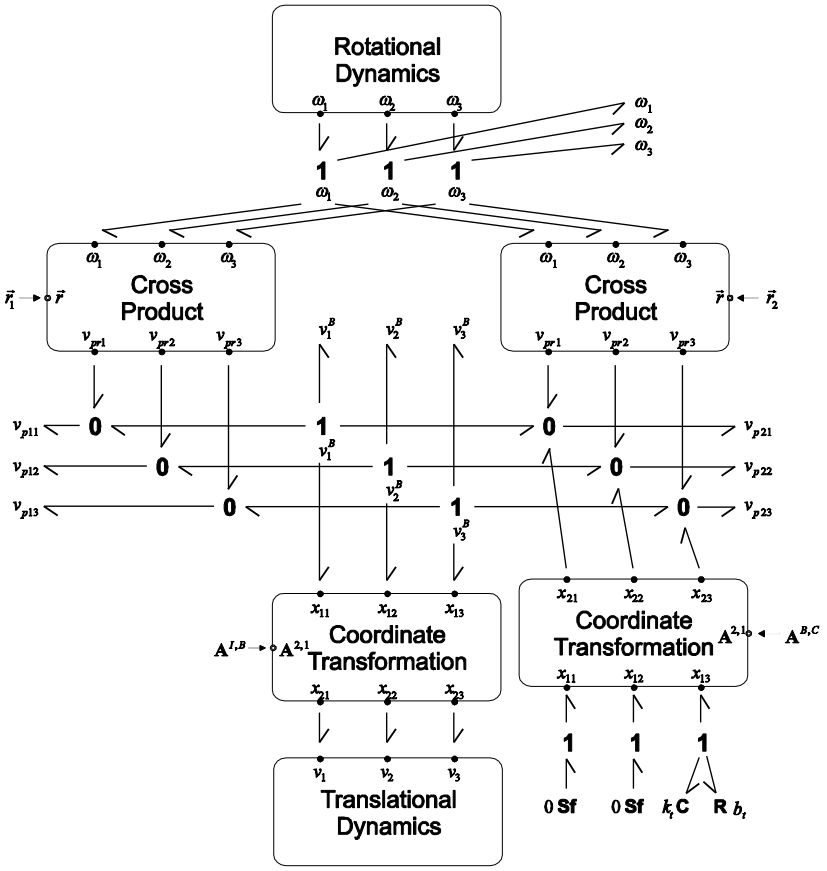


Figure 6.7. The multibody model of the tire

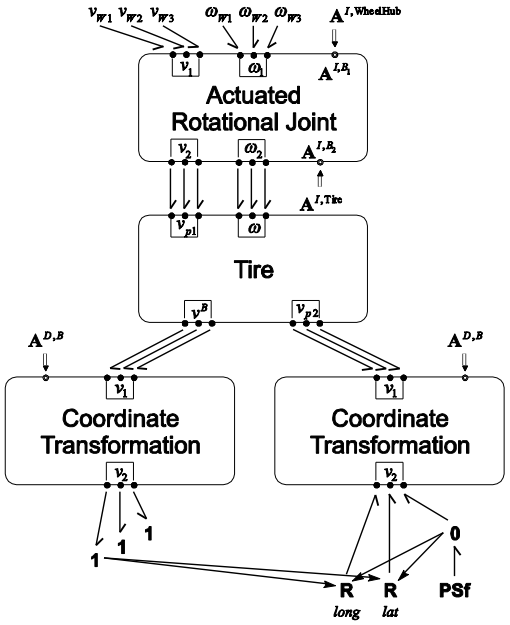


Figure 6.8. The tire and road model

The tire slip models are adopted from DADS [236] as follows. The longitudinal force is given by

$$F_x = \begin{cases} F_z \mu_f & \text{acceleration} \\ -F_z \mu_f & \text{breaking} \end{cases} \tag{6.1}$$

where F_z is the normal force, and μ_f is given by the piecewise linear function of slip, λ_x,

$$\mu_f = \begin{cases} \mu \lambda_x / 0.2 & \lambda_x < 0.2 \\ \mu(\lambda_x + 0.8) & 0.2 \le \lambda_x < 0.25 \\ \mu(-0.2\lambda_x + 1.1) & 0.25 \le \lambda_x < 0.5 \\ \mu & \lambda_x \ge 0.5 \end{cases} \tag{6.2}$$

with μ being the friction coefficient, and λ being defined as

$$\lambda_x = |\kappa| = \left| \frac{\omega R - v_x}{v_x} \right| \tag{6.3}$$

where ω, R, and v_x are the rotational speed, radius, and longitudinal velocity of the tire. The sketch of the $F_x - \kappa$ curve is shown in Fig. 6.9.

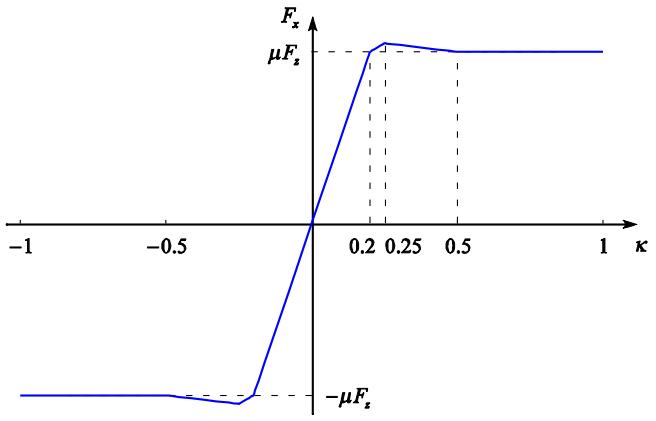


Figure 6.9. The $F_x - \kappa$ curve

The lateral tire force, on the other hand, is given by

$$F_y = \begin{cases} -F_{lat} & v_y > 0 \\ F_{lat} & v_y \le 0 \end{cases} \tag{6.4}$$

where F_{lat} is given by

$$F_{lat} = \begin{cases} a_1\lambda_y^3 + a_2\lambda_y^2 + a_3\lambda_y & \lambda_y < \alpha_n \\ \mu F_z & \lambda_y \geq \alpha_n \end{cases}$$
$$a_1 = \frac{-2\mu F_z + C_a\alpha_n}{\alpha_n^3};\ a_2 = \frac{3\mu F_z - 2C_a\alpha_n}{\alpha_n^2};\ a_3 = C_a \tag{6.5}$$

The coefficients a_i are chosen such that the slope at the origin is the cornering stiffness, C_a, and the slope is zero and F_{lat} is equal to μF_z when $\alpha = \alpha_n$. α_n and the lateral slip , λ_y, are defined as

$$\alpha_n = 2.5\frac{F_z}{C_a}$$
$$\lambda_y = |\alpha| = \left|\tan^{-1}\left(\frac{v_y}{v_x}\right)\right| \tag{6.6}$$

v_y is the lateral speed of the tire. The sketch of the $F_y - \alpha$ curve is shown in Fig. 6.10.

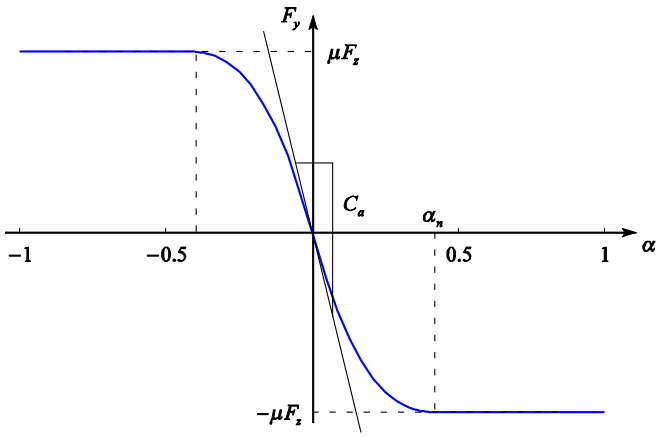


Figure 6.10. The $F_y - \alpha$ curve

6.2.4. The Steering Mechanism

The steering mechanism consists of the idler and Pitman arms, the steering link, and the two tie rods connecting the mechanism to the wheel hubs. The mechanism is driven through the Pitman arm. The steering mechanism is shown in Fig. 6.11, and its model is given in Fig. 6.12.

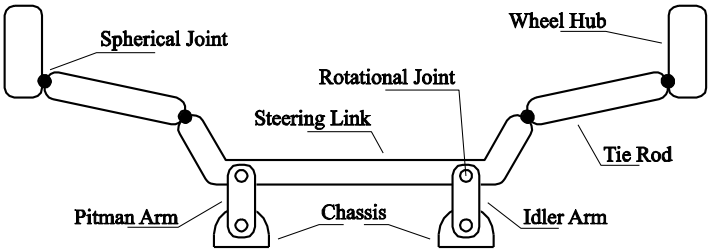


Figure 6.11. The steering mechanism

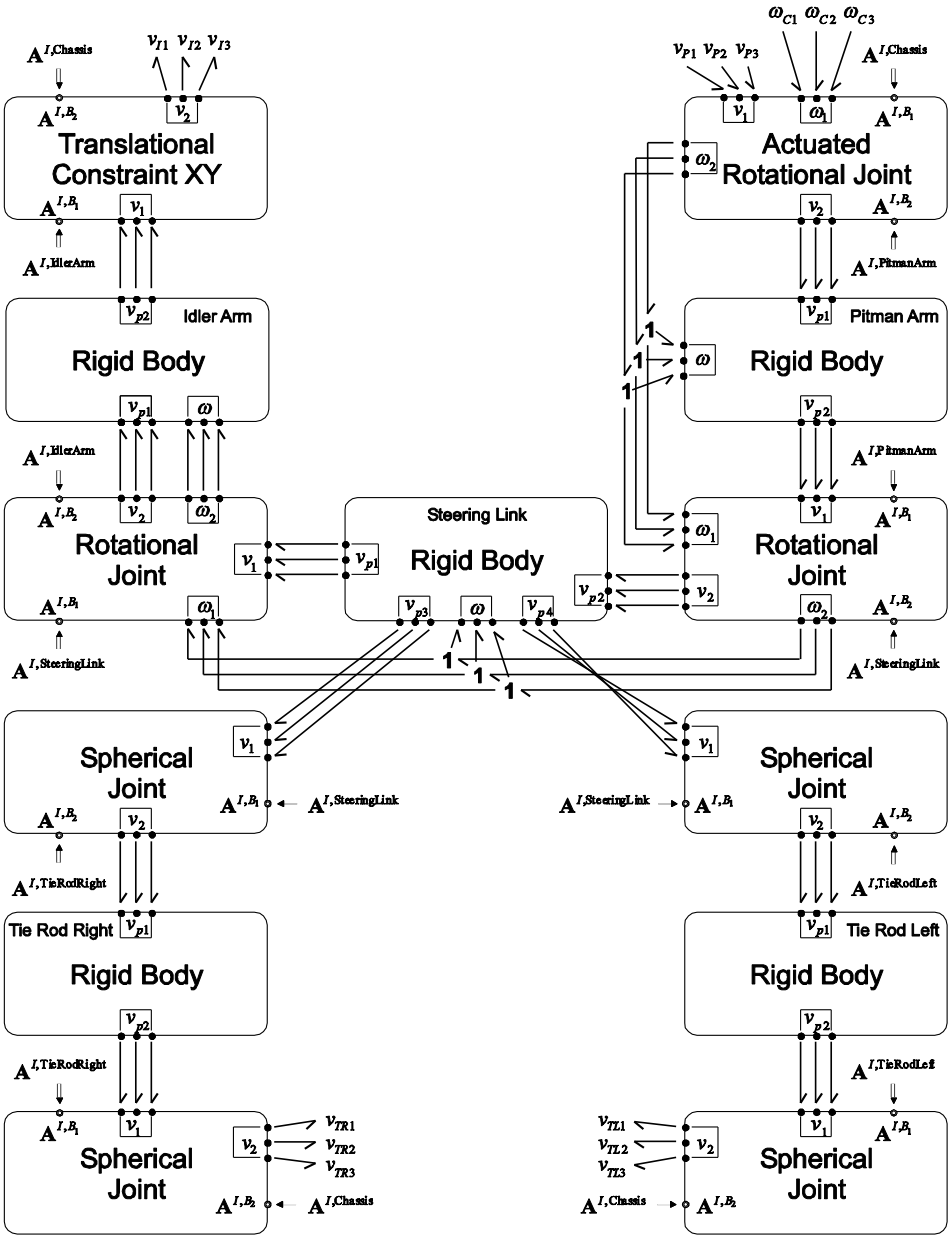


Figure 6.12. The model of the steering mechanism

6.2.5. Assembling the HMMWV Model

Once the subcomponent models are obtained as above, the HMMWV model can be assembled as shown in Fig. 6.13. In addition to the components discussed so far, the assembled model also includes the effects of the front and rear anti-roll bars implemented as spring elements applying restoring forces on the velocity differences between the left and right suspension struts. This model will hereafter be referred to as the full model.

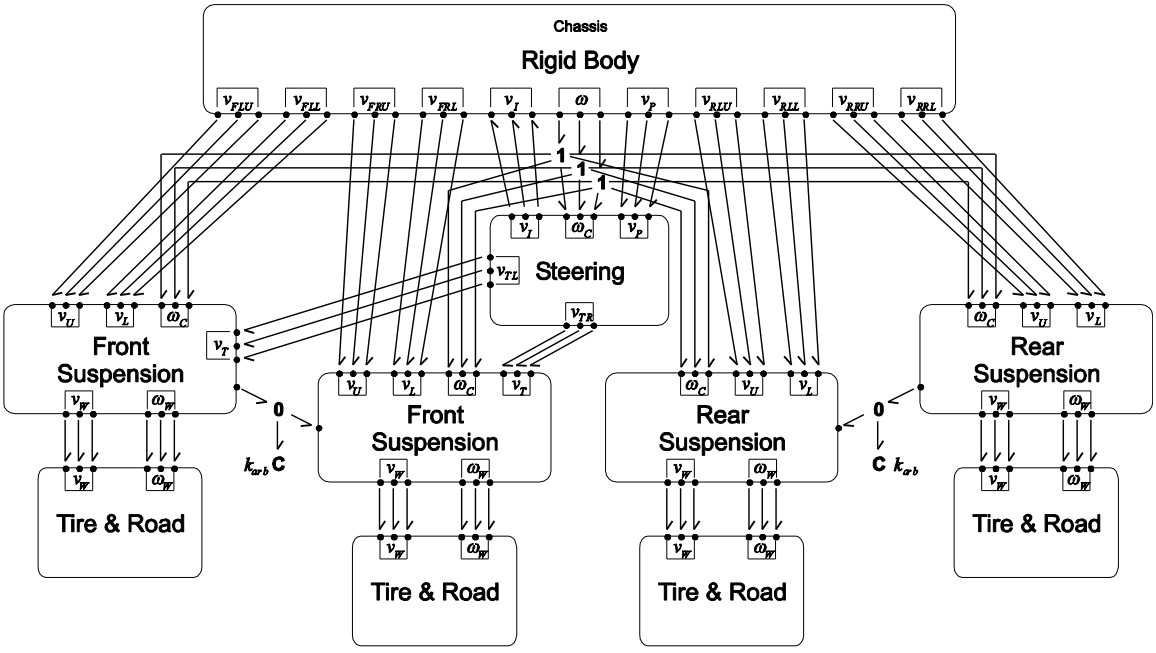
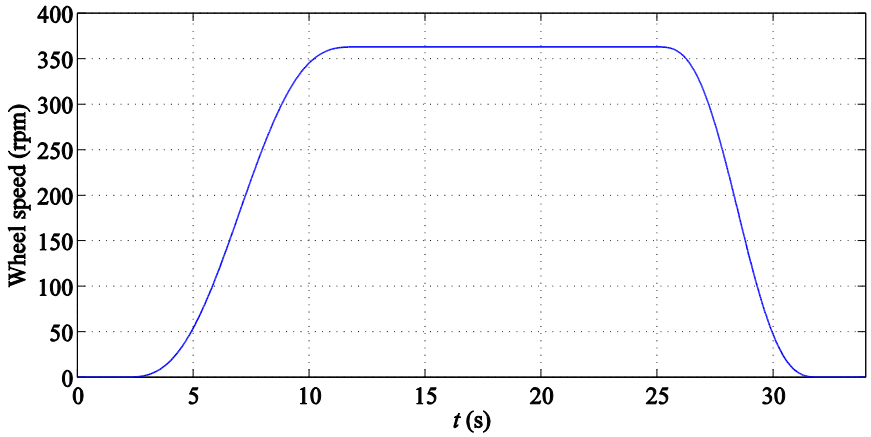


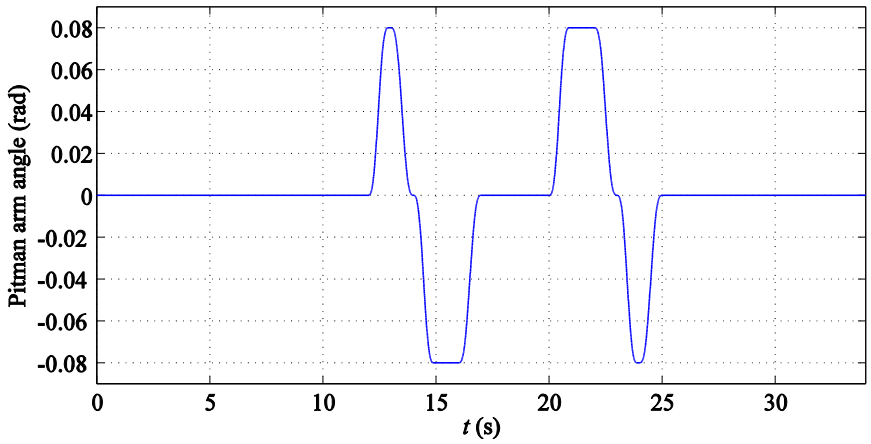
Figure 6.13. The full HMMWV model

6.3. SCENARIOS OF INTEREST

The first scenario of interest is a two-double-lane-change maneuver on a flat road. Such a maneuver could be employed in, e.g., vehicle rollover and handling studies. The particular inputs considered for this scenario are given in Fig. 6.14. The output of interest is the roll acceleration of the vehicle.



(a)



(b)

Figure 6.14. The inputs used in the lane-change scenario: (a) the wheel speed; (b) the Pitman arm angle

The second scenario is the shaker table scenario. This scenario might be of interest when studying, e.g., the suspension characteristics and ride-quality of a vehicle. In this scenario the tires are removed from the model, and a sinusoidal sweeping displacement with an amplitude of 5 cm and frequency range of 0-8 Hz is applied to all four wheel hubs. Figure 6.15 shows the particular input considered. The output of interest is the z-position of the chassis.

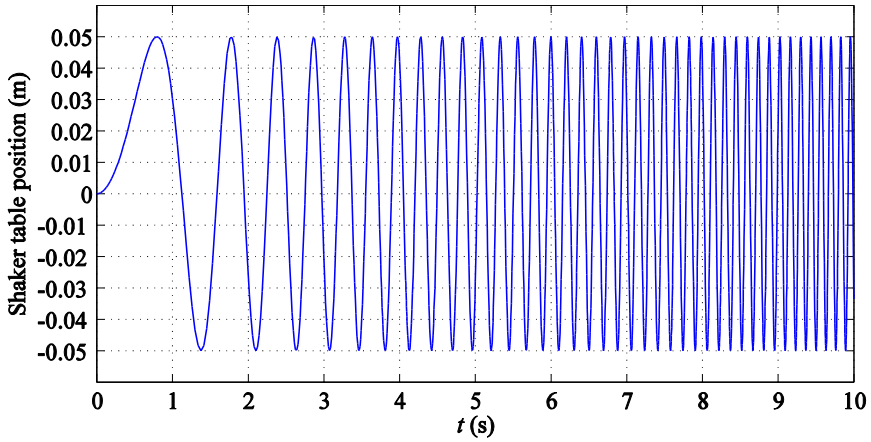


Figure 6.15. The shaker table displacement input used for the shaker table scenario

The third scenario is driving straight on a flat road. This scenario could be useful, e.g., when studying the acceleration characteristics of the vehicle, sizing the engine, or designing a cruise controller. In this particular scenario the input dictates the wheel speed, as shown in Fig. 3.1. The output of interest is the longitudinal acceleration of the vehicle.

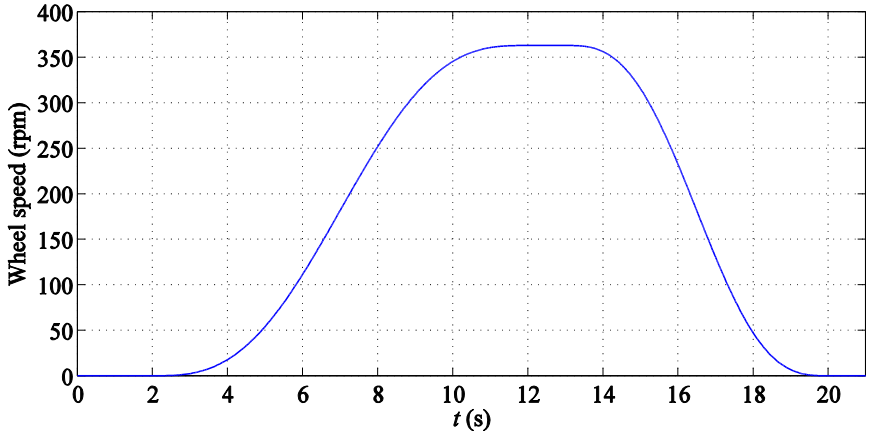


Figure 6.16. The wheel speed input used in the straight-driving scenario

6.4. SIMPLIFICATION AND REDUCTION OF THE FULL HMMWV MODEL

As shown above, the modular approach enables the rapid modeling of systems as complex as the HMMWV. However, the resulting model may not be proper for certain

scenarios of interest. Indeed, the full model obtained above is overly complex for all the three scenarios considered. Therefore, the methods developed throughout this work are employed here to obtain proper models for the scenarios of interest.

First, the full HMMWV model is simplified using the algorithm proposed in Chapter 3. It is desired to obtain one simplified model that is valid for all three scenarios. Thus, this exercise also demonstrates how multiple scenarios can be taken into account. To this end, the full model is simulated and the junction inactivity analysis is performed for the three scenarios of interest. Each scenario produces a different set of inactive junctions. The intersection of these three sets is then taken, and the full model is simplified using this set. Table 6.1 highlights what is removed from the model during simplification. There are other structural simplifications in the model in addition to what is shown in Table 6.1, especially in cross products and coordinate transformations, but due to space considerations the details of these have been omitted.

Table 6.1. Highlights of simplified parts of the full model

	Translational		Rotational	
	Dynamics	Kinematics	Dynamics	Kinematics
Front Suspensions				
Rot. Joints	–		–	x
Upper A-Arm			y[1], z	
Rear Suspensions				
Rot. Joints	–		–	x
Upper A-Arm			y, z	
Steering				
Pitman Arm			x, y, z	x
Trans. Con. XY	–	z	–	–
Rot. Joints	–		–	z
Steering Link			y	
Idler Arm			x, y, z	x
Tie Rods			y	y

[1] Rotational dynamics in y-axis is kept in the model in the front left suspension.

The simplified model is then reduced for each scenario separately using the algorithm proposed in Chapter 5. For the two-double-lane-change scenario the model is first reduced using the global application of the proposed metric. Further reduction is then

obtained by the local application of the metric to the energy storing/dissipating elements only. For the other scenarios, only the global approach is used. Figures 6.17-6.19 show the global reduction thresholds obtained in each scenario. The thresholds below a relative importance of 10^{-9} are not shown, and the bond number in bold indicates the threshold chosen for reduction. The thresholds are chosen to achieve highest levels of reduction with an acceptable compromise in accuracy.

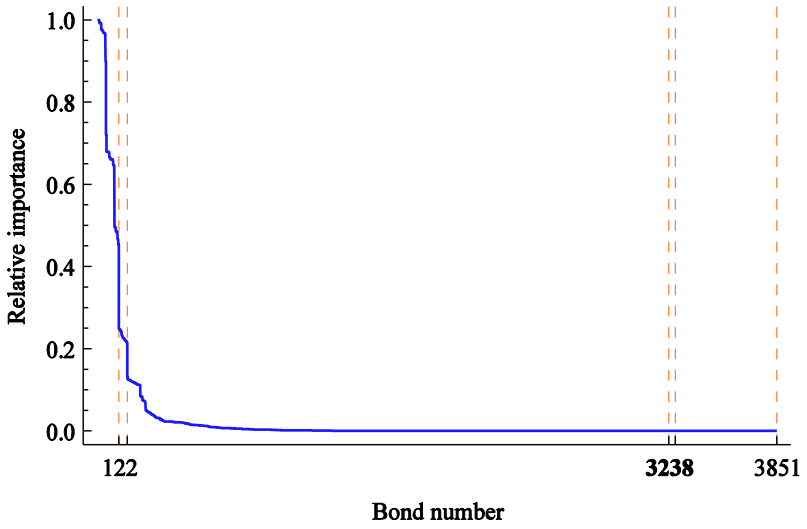


Figure 6.17. Thresholds for $r = 1.5$ in the lane change scenario

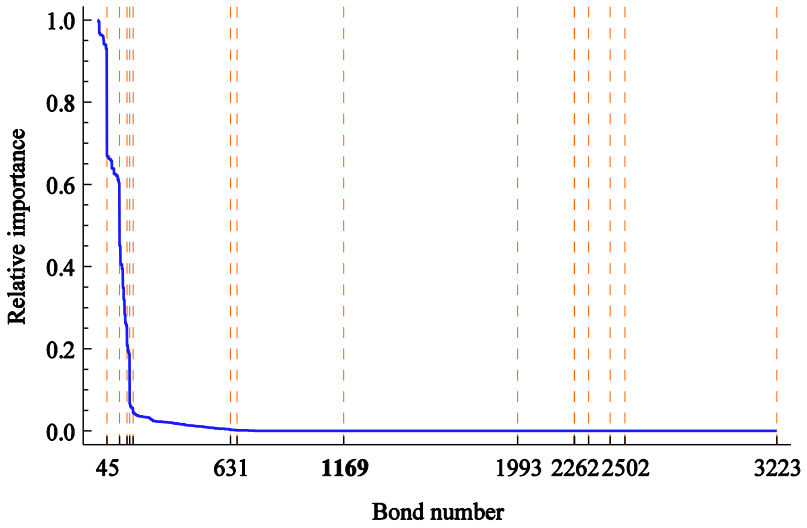


Figure 6.18. Thresholds for $r = 1.2$ in the shaker table scenario

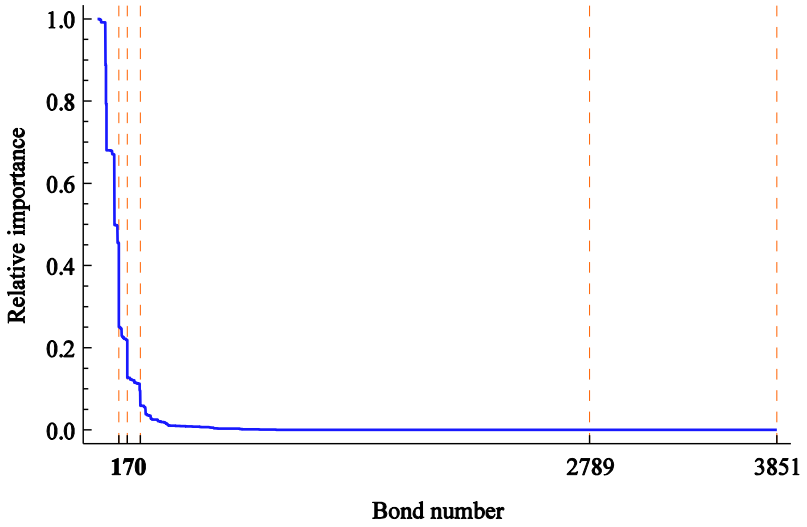


Figure 6.19. Thresholds for $r = 1.5$ in the straight driving scenario

Tables 6.2- 6.4 highlight what is removed from the model in each scenario. Parts that were already removed during simplification are shown in grey. Once again, due to space considerations, these Tables do not give an exhaustive list.

Table 6.2. Highlights of reduction for the lane change scenario

	Translational		Rotational	
	Dynamics	Kinematics	Dynamics	Kinematics
Front Suspensions				
Rot. Joints			–	x
Upper A-Arm	z		x, y[1], z	
Wheel Hub	z[2]		x, y, z	
Lower A-Arm	z		x, y, z	
Rear Suspensions				
Rot. Joints			–	x
Upper A-Arm	z		x, y, z	
Wheel Hub	z[2]		x, y, z	
Lower A-Arm	z		x, y, z	
Tires & Roads				
Wheel	z		x[2], z[2]	
Steering				
Pitman Arm	z		x, y, z	x
Trans. Con. XY	–	z	–	–
Rot. Joints	–		–	z
Steering Link	z		x, y, z	
Idler Arm	z		x, y, z	x
Tie Rods	z		x, y, z	y

[1] In the right suspension this was already removed during simplification.
[2] These are removed by the local analysis.

Table 6.3. Highlights of reduction for the shaker table scenario

	Translational		Rotational	
	Dynamics	Kinematics	Dynamics	Kinematics
Chassis	x, z	y	x, y, z	x, z
Front Suspensions				
Rot. Joints			–	x, y, z
Upper A-Arm	x, y, z		x, y[1], z	y, z
Wheel Hub	x, y		x, y, z	
Lower A-Arm	x, y		x, y, z	y, z
Rear Suspensions				
Rot. Joints	x[2]		–	x, y, z
Tans. Con. YZ	–	x	–	
Upper A-Arm	x, y, z		x, y, z	y, z
Wheel Hub	x, y		x, y, z	y, z
Lower A-Arm	x, y	x	x, y, z	y, z
Steering				
Act. Rot. Joint	–	y	–	x, y, z
Pitman Arm	x, y, z	y	x, y, z	x, y, z
Trans. Con. XY	–	y, z	–	–
Rot. Joints	–	y, z[3]	–	x, y, z
Steering Link	x, y, z	y	x, y, z	x, y, z
Idler Arm	x, y, z	y, z	x, y, z	x, y, z
Tie Rods	x, y, z	x	x, y, z	y, z

[1] In the right suspension this was already removed during simplification.
[2] Only in the rotational joint between wheel hub and lower arm
[3] Only in the rotational joint between steering link and idler arm

Table 6.4. Highlights of reduction for the straight-driving scenario

	Translational		Rotational	
	Dynamics	Kinematics	Dynamics	Kinematics
Chassis	y, z	y, z	x, y, z	x, y, z
Front Suspensions				
Rot. Joints	–	y, z	–	x, y, z
Spherical Joints	–	y, z	–	–
Upper A-Arm	x, y, z	y, z	x, y[1], z	x, y, z
Wheel Hub	x, y, z	y, z	x, y, z	x, y, z
Lower A-Arm	x, y, z	y, z	x, y, z	x, y, z
Rear Suspensions	entirely removed			
Front Tires&Roads				
Rot. Joint	–	y	–	x, z
Wheel	x, y, z	y	x, y, z	x, z
Long. Slip	removed			
Lat. Slip	removed			
Rear Tires&Roads	entirely removed			
Steering	entirely removed			

[1] In the right suspension this was already removed during simplification.

It is worth noting that in the last scenario the model essentially reduces down to two disks rolling without slip.

Tables 6.5-6.7 compare the full, simplified and reduced models in various aspects for the three scenarios. These Tables highlight the computational efficiency gained by using the proposed algorithms. Recall that the simplified model is obtained for all three scenarios. Better results would be obtained with simplification, if the full model was simplified for each scenario separately.

Table 6.5. Comparison of the models for the lane-change scenario

	Full	Simplified		Reduced	
		quantity	% decrease	quantity	% decrease
Processing time [min]	8.0	4.7	41.0%	4.5	43.8%
Number of equations	8126	6034	25.7%	4967	38.9%
Number of variables	10263	7657	25.4%	6330	38.3%
Number of independent states	226	199	11.9%	176	22.1%
Number of dependent states	72	83	-15.3%	44	38.9%
Number of constraints	50	27	46.0%	34	32.0%
Simulation time [s]	138.5	86.4	37.6%	16.3	88.2%
Simulation speed	×1	×1.6		×8.5	

Table 6.6. Comparison of the models for the shaker table scenario

	Full	Simplified		Reduced	
		quantity	% decrease	quantity	% decrease
Processing time [min]	4.6	2.9	37.4%	0.5	88.5%
Number of equations	6333	4729	25.3%	1824	71.2%
Number of variables	8006	5888	26.5%	2694	66.4%
Number of independent states	186	159	14.5%	35	81.2%
Number of dependent states	54	65	-20.4%	8	85.2%
Number of constraints	48	25	47.9%	40	16.7%
Simulation time [s]	17.7	14.0	20.6%	0.9	94.7%
Simulation speed	×1	×1.3		×18.8	

Table 6.7. Comparison of the models for the straight-driving scenario

	Full	Simplified		Reduced	
		quantity	% decrease	quantity	% decrease
Processing time [min]	8.2	4.7	43.2%	0.01	99.9%
Number of equations	8054	5962	26.0%	152	98.1%
Number of variables	10232	7626	25.5%	234	97.7%
Number of independent states	226	199	11.9%	12	94.7%
Number of dependent states	72	83	-15.3%	0	100%
Number of constraints	50	27	46.0%	0	100%
Simulation time [s]	48.5	34.3	29.2%	0.1	99.8%
Simulation speed	×1	×1.4		×437	

Finally, Fig. 6.20-6.22 show the outputs of the full, simplified and reduced models for the three scenarios. The approximation errors are given in Table 6.8.

Table 6.8. Approximation errors of the simplified and reduced models

	Maximum error		Relative L_2-norm of error	
	Simplified	Reduced	Simplified	Reduced
Lane change	1.85144×10^{-4}	2.41157×10^{-3}	3.72147×10^{-4}	6.7397×10^{-3}
Shaker table	4.98328×10^{-6}	1.64458×10^{-3}	3.4559×10^{-5}	5.54002×10^{-2}
Straight driving	8.00417×10^{-3}	0.755391	6.63184×10^{-5}	0.1348

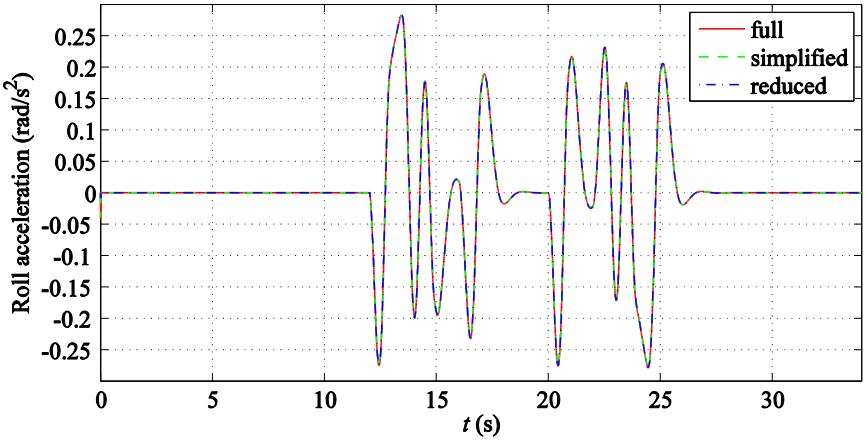


Figure 6.20. Outputs of the full, simplified and reduced models for the lane-change scenario

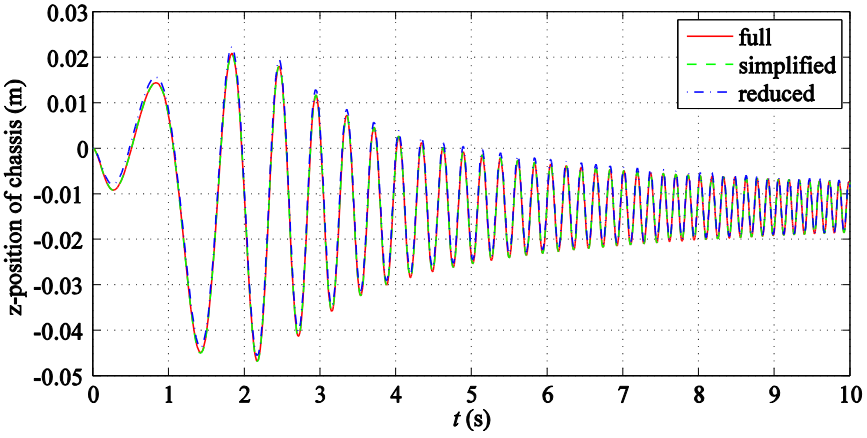


Figure 6.21. Outputs of the full, simplified and reduced models for the shaker table scenario

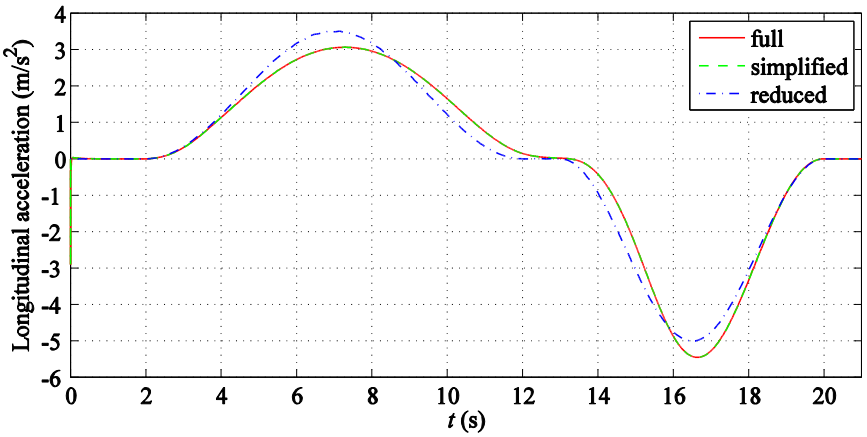


Figure 6.22. Outputs of the full, simplified and reduced models for the straight-driving scenario

6.5. DISCUSSION OF THE RESULTS

The first part of this case study verifies the benefits of the modular approach to modeling, especially for complex systems such as the HMMWV. The subcomponents are modeled independently, and many model blocks, such as rigid bodies, joints, and coordinate transformations, have been reused during modeling. The submodels are verified independently, and the full model is assembled rapidly from the component models.

The second part of the case study analyzes the performances of the proposed algorithms. First of all, both the simplification and reduction algorithms are successfully applied to the rather large-scale nonlinear full model. The scenarios of interest are explicitly taken into account during the simplifications and reductions. The simplified and reduced models preserve the realization, and thus the physical meaning of the full model. Simplifications and reductions are achieved not only in the dynamics, but also in the kinematics. Finally, simplifications and reductions are obtained at the graph-level, preserving the intuitive appeal of the graph representation.

Note that, as expected, the outputs of the simplified model accurately follow the outputs of the full model in all three scenarios with less computational cost. Furthermore,

with some user-controllable compromise in accuracy, the reduced models can improve the conceptual and computational efficiency of the simplified model even further.

It is important to keep in mind that the reduced models presented above are results of particular choices of reduction thresholds. In this case study the thresholds are chosen as high as possible to achieve maximum reduction. If a higher fidelity is desired, it is possible to work with lower thresholds.

The lane-change and shaker table scenarios make it clear that the proposed importance metric would benefit from being output-dependent as well. In both cases there are model parts that can be eliminated without compromising the output accuracy too much, however since they are within the same threshold as other parts that are important to keep in the model, they are kept in the model as well. For example, in the shaker table scenario the steering system can be safely eliminated from the model, but that also means the elimination of the vertical kinematics of the upper arms at the points where they are connected to the chassis, which are critical to keep in the model. Similarly, in the lane-change scenario the lateral dynamics of the Pitman arm could be removed, but then the anti-roll bars would be removed as well, because they are within the same threshold. However, the anti-roll bars are clearly important in a rollover study and should not be eliminated.

These cases exemplify the persisting challenge that some relatively high-energy components might be kept in the model by the proposed algorithm even though they might be unimportant for certain outputs. In that case, local application of the metric could be one option for overcoming this challenge, i.e., one can perform the analysis with a subset of the graph that leaves out the parts that are to be kept in the model. This is indeed what has been done in the lane-change scenario. The global application of the analysis was followed by the local analysis with the energy storing/dissipating elements to further remove some dynamics that were not removed by the global analysis because of some important kinematics that were within the same threshold. Nevertheless, the local

analysis offers only a limited solution to the problem, since it is not always clear to see which bonds isolate. A sensitivity analysis reveals that a better solution can be obtained by adding output dependence. Specifically, the sensitivity of the output (as measured by the integral-of-squared-value criterion) to a 10% change in the stiffnesses of the front and rear anti-roll bars is 13% and 32%, respectively, whereas it is only 0.02% for a 10% change in the Pitman arm mass. Thus, it is beneficial to find an efficient way to incorporate the outputs of interest in the analysis. This is left as future work.

The sensitivity results reported above are obtained with the full model. When the same analysis is performed using the reduced model, the output sensitivity to front and rear anti-roll bar stiffnesses and the Pitman arm mass are obtained as 14%, 32%, and 0.4%, respectively. The close agreement between the full and reduced models indicates that even though the model was reduced for a specific set of parameters, it is still valid for the considered change in parameters. However, a formal study of the range of validity of the models reduced with the proposed method is left as future work.

This case study also exposed two computational issues with the proposed reduction method, which may pose not fundamental, but practical limitations for large-scale systems. First, the singular value decomposition associated with the proposed metric can become computationally expensive when the number of bonds and/or the number of observations increase (although still not as expensive as performing a full-scale sensitivity analysis). Second, if the energy calculations are performed at each bond during the simulation, the simulation can slow down significantly due to the increase in the number of states. Applying the simplification algorithm first before performing reduction alleviates both issues by decreasing the number of bonds. Furthermore, the energy observation matrix may be resampled to decrease the number of observations. Both approaches have been used in this case study to make the reduction problem more numerically tractable. In addition, the energy calculations may be performed offline to

avoid an increase in the number of states, but this needs further investigation as to whether aliasing would be an issue or not.

6.6. SUMMARY AND CONCLUSION

The proper modeling of a HMMWV for three different scenarios is presented as a case study. The multibody dynamics of the vehicle is first modeled modularly to obtain the full model. The full model is then simplified and reduced for the following scenarios: two double-lane-change maneuvers, shaker table, and driving straight. The full, simplified and reduced models are compared in various aspects.

The conclusion is that the proposed algorithms can successfully simplify and reduce even very complex systems. Moreover, the modeling philosophy adopted in this work – modular modeling accompanied by simplification/reduction – has been shown to be an efficient way for proper modeling of complex systems. This case study also verifies that incorporating output-dependence to the proposed methods is an important future work, if the properness of models is to be maximized for the considered outputs.

CHAPTER 7

CONCLUSION

"*Simplicity is the ultimate sophistication*."
– Leonardo da Vinci

7.1. SUMMARY

Proper models are critical for efficient simulation and design. They are defined as models that balance fidelity and simplicity, or accuracy and complexity. These two goals are typically contradictory. Moreover, achieving them simultaneously requires much time and expertise. Therefore, obtaining proper models is challenging.

This work has been done to advance the knowledge in the domain of proper modeling of dynamic systems. Specifically, a review of literature revealed that there is a need for methods that can handle nonlinear models at graph-level, are realization-preserving and trajectory-dependent, and target not only the order, but also the structure of a model.

With a reductive approach to proper modeling, i.e., assuming that a model with satisfactory accuracy and excessive complexity exists, this work first set out to simplify the structure of models. Models obtained through a modular approach have been considered, because these models are typically prone to having an excessive complexity in them. The concept of junction-inactivity has been introduced, and a simplification algorithm based on this concept has been developed. Examples are used to illustrate the mechanics and benefits of the algorithm, as well as to discuss its important properties.

One of these properties was preserving the realization of a given model. While this property was listed as one of the desired properties of the algorithms sought in this work, the work also explored the possibility of allowing a change in realization to a limited extent. Particularly, the orientation of the body-fixed coordinate frames in multibody systems has been considered. The rationale for this focus was that reorienting a coordinate frame would not change the physical meaning of the coordinates, which was the main goal behind preserving the realization in the first place. With this in mind, an algorithm has been proposed that can automatically reorient body-fixed coordinate frames in a multibody system to render the realization more conducive to simplification. An example illustrated that the reorientation algorithm complements the simplification algorithm well.

The work then moved to the reduction of models beyond just simplification. To this end, a new energy-based metric has been proposed to assess the relative importance of dynamic and structural components in a system to the system's overall behavior. Specifically, the metric applied the Karhunen-Loève expansion to energy trajectories within the system to identify dominant energy flow patterns in a given realization. Based on this metric a reduction algorithm has been proposed that can yield simultaneous order and structure reduction. Examples demonstrated the mechanics and advantages of the proposed algorithm.

Finally, a case study has been done with the High-Mobility-Multipurpose-Wheeled-Vehicle (HMMWV), a system that is representative of the contemporary engineering systems. The multibody dynamics of the vehicle have been modeled by a modular approach. This full model has then been simplified and reduced for three different scenarios through successful applications of the developed algorithms.

7.2. LIST OF CONTRIBUTIONS

This work contributes to the proper modeling literature in the following aspects:

1. The inactive-junction concept has been introduced. This concept builds on the existing notion of activity and extends it to junction elements. Furthermore, this concept generalizes the well-known ideas that zero-flow 1-junctions and zero-effort 0-junctions can be removed from a bond graph model to simplify it.

2. A structural simplification algorithm has been developed based on the inactive-junction concept. This algorithm can yield simplifications beyond what can be achieved with existing rules and methods. Specifically, it yields trajectory-dependent realization-preserving simplifications in nonlinear models at graph-level, and its numerical nature makes it able to handle not only exact, but also approximate cases. For example, the algorithm can simplify not only ideal constraints, but also constraints that are approximated using dynamic elements such as springs and dampers.

3. A Karhunen-Loève-expansion-based algorithm has been developed to reorient body-fixed coordinate frames in multibody systems and render the realization more conducive to simplification. Due to this very specific focus, unlike some existing methods, the physical meaning of the coordinates is preserved.

4. An energy-based metric has been developed to assess the relative importance of dynamic and structural components in a system. This metric, for the first time, combines the correlation-based idea behind the Karhunen-Loève expansion with the energy-based reduction philosophy. This uniquely allows

for the unified assessment of dynamic and structural components, or, in other words, dynamics and kinematics.

5. A reduction algorithm based on this metric has been developed that can reduce not only the order, but also the structure of a model. This algorithm is also trajectory-dependent, realization-preserving, and can handle graph-level representations of nonlinear systems.

6. Proper models of the multibody dynamics of a HMMWV have been created for three different scenarios. This is achieved by first developing a very detailed multibody model of the HMMWV using the modular approach, and then simplifying and reducing the model using the proposed algorithms.

7.3. FUTURE WORK

As pointed out throughout this document, this work presents many possible directions for future work. These can be summarized as follows:

1. All algorithms developed in this work are trajectory dependent. Choosing the input trajectory that properly describes the scenario of interest was left to the modeler in this work. However, this may itself be a challenging task. For example, one may legitimately seek a model that is proper for a set of various excitations, or a range of parameter values. One possible way of handling the first case was presented during the simplification of the HMMWV model, where the junction-inactivity-analysis results of the three different scenarios were combined to obtain one simplified model valid for all three scenarios. This approach, however, may not always be feasible, especially if a large or infinite number of input trajectories or parameter values are considered, as is the case with a continuous range of design parameters. Can one then use a

feasible number of sampled inputs/parameters? If yes, how to choose the samples? Or, is it possible to create a single simulation run to cover all inputs or parameter values of interest? If yes, how? These are important questions that need to be answered.

2. Instead of repeating the analyses for several trajectories, it may also be possible to determine the range of validity of the simplified and reduced models obtained with the proposed methods by utilizing a Lyapunov-function-type approach. Such an analysis would be beneficial for design and control.

3. Both the simplification and reduction algorithms developed in this work are realization preserving. The main motivation behind this property is to preserve the physical meaning of the original realization. The coordinate-frame-reorientation algorithm, however, investigated one possible way of changing the realization without changing the physical meaning to make the model more conducive to simplification. There may be other possibilities, and this could be investigated further.

4. The coordinate-frame-reorientation algorithm was created for the specific rigid-body representation considered throughout this work. Particularly, the rotational dynamics were expressed in a body-fixed frame located at the center of mass, and the translational dynamics were expressed in the inertial frame. Although this is a widely-used representation, the algorithm could be extended to more general representations.

5. The proposed algorithms are not output dependent. This means that some model parts may not be eliminated even if they do not contribute to the outputs of interest, because they may be contributing to some other outputs.

To take most advantage of simplification/reduction, output dependence should also be incorporated into the proposed algorithms.

6. The importance metric introduced in this work was utilized for reduction purposes only. However, it may also be useful in other applications such as partitioning, scaling, or system identification. Particularly, the literature shows successful applications of the activity metric to partitioning [19, 201], scaling [237] and system identification [216], and the proposed metric may improve those results. This should be investigated further.

7. It may also be interesting to investigate the possibility, advantages and disadvantages of implementing the proposed algorithms as part of a numerical solver. It may be possible to increase the computational efficiency of numerical solvers by performing simplifications/reductions online as the integration proceeds.

8. Only deterministic systems were considered in this work. Extension to stochastic systems is another important future direction. How the proposed methods can be used or should be extended to identify which uncertainties are important to model and which are negligible is a very interesting question.

9. Only continuous models were considered. There are many engineering systems, however, that would be best modeled using hybrid models, like the clutch in a transmission. Proper modeling of hybrid systems is also an important challenge, and the proposed algorithms should be extended to hybrid systems as well.

10. This work concentrated on energetic systems. As such, the methods developed cannot be readily applied to systems in which there is no energy, but

information flow, such as control systems or economic systems. If proper modeling of such systems is of interest, the proposed methods should be extended to be able to handle signals as well.

7.4. CONCLUSION

The work presented in this document supports the two hypotheses stated in Chapter 1. Specifically, the activity metric has been successfully utilized as a structural simplification tool; and the importance metric has been shown to give a good assessment of various dynamic and structural components of model, thereby helping reduce not only the order, but also the structure. The identified need for realization-preserving trajectory-dependent structure and order simplification/reduction algorithms that are applicable to nonlinear systems at the graph level has been partially addressed, and opportunities for further research have been pointed out.

APPENDICES

APPENDIX A

BOND GRAPH MODEL LIBRARY

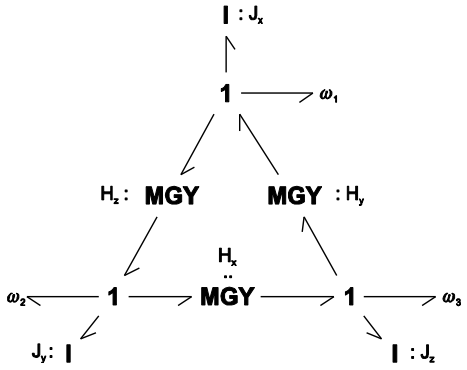


Figure A.1. Rotational dynamics

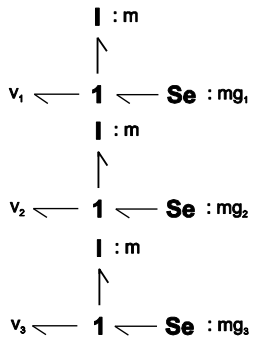


Figure A.2. Translational dynamics

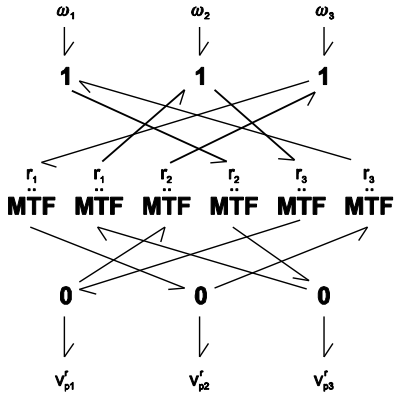


Figure A.3. Cross product

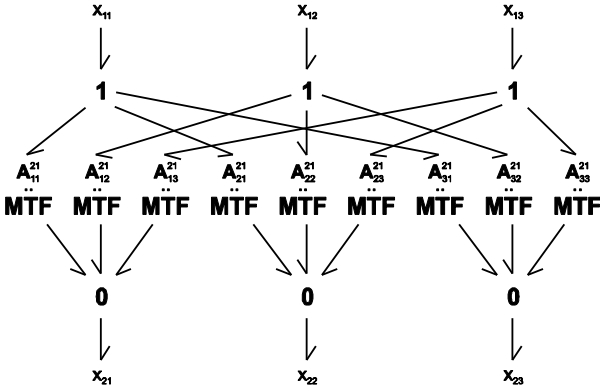


Figure A.4. Coordinate transformation for flow-in-flow-out causality

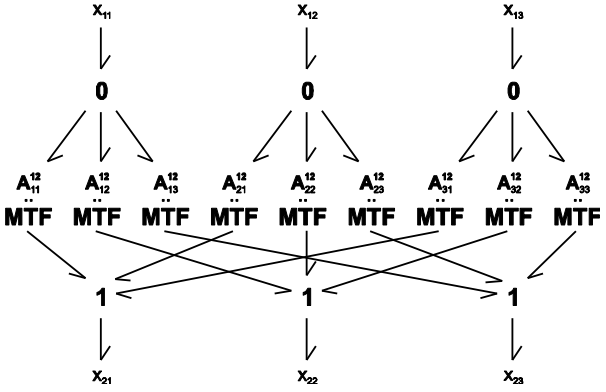


Figure A.5. Coordinate transformation for effort-in-effort-out-causality

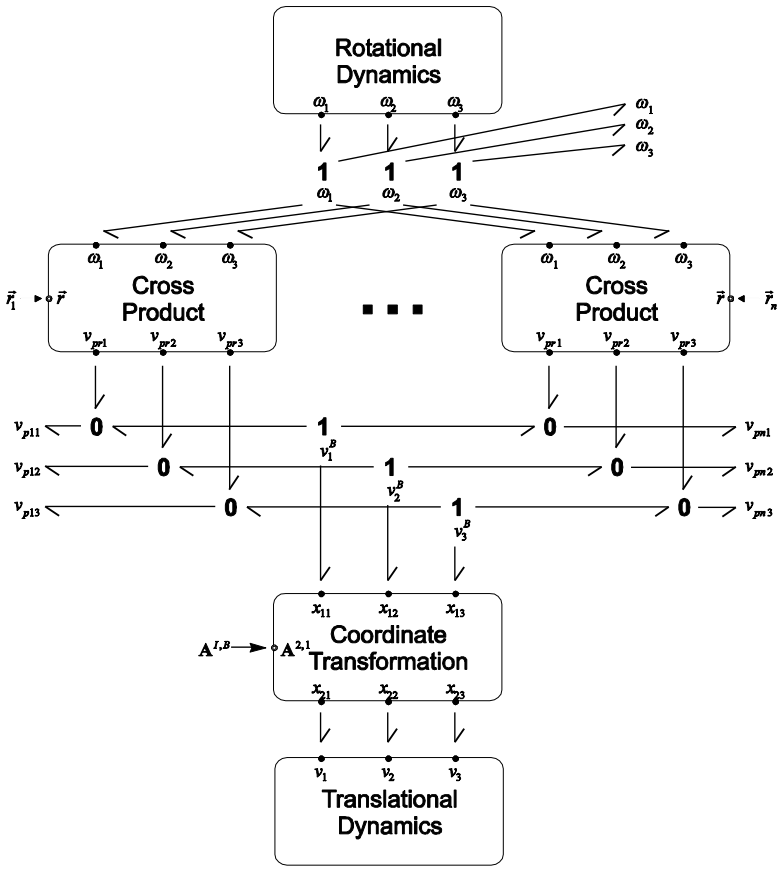


Figure A.6. Rigid body

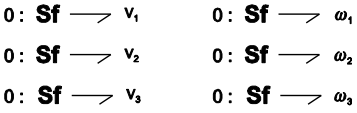


Figure A.7. Ground

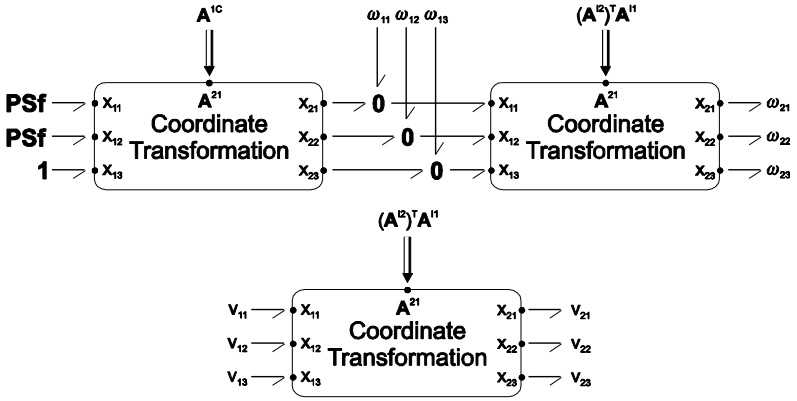


Figure A.8. Rotational joint for flow-in-flow-out causality in translational domain

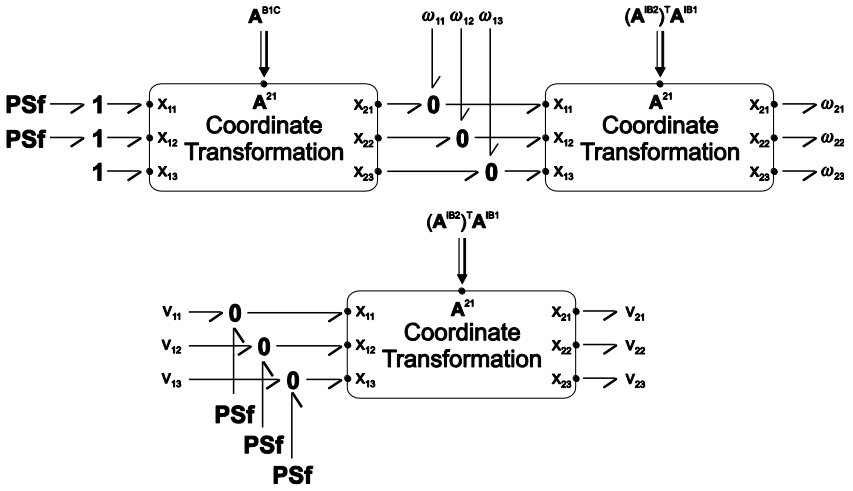


Figure A.9. Rotational joint for effort-in-effort-out causality in translational domain

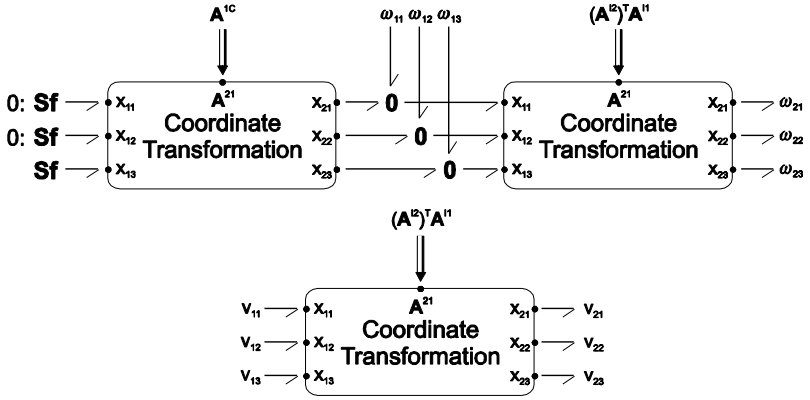


Figure A.10. Actuated rotational joint

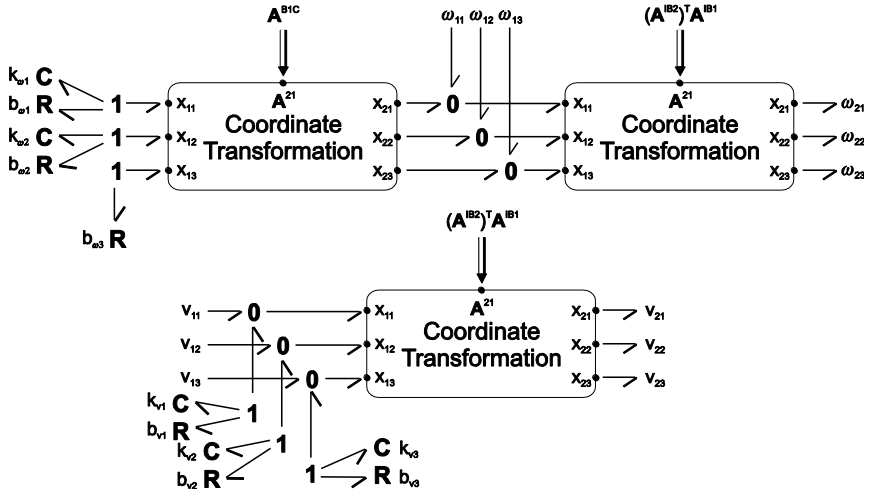


Figure A.11. Real rotational joint

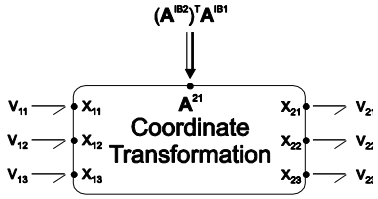


Figure A.12. Spherical joint for flow-in-flow-out causality

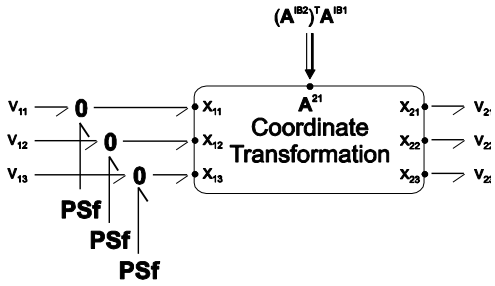


Figure A.13. Spherical joint for effort-in-effort-out causality

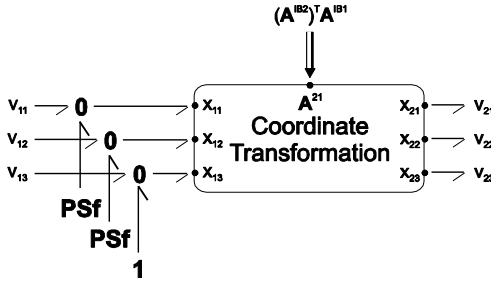


Figure A.14. Translational constraint XY

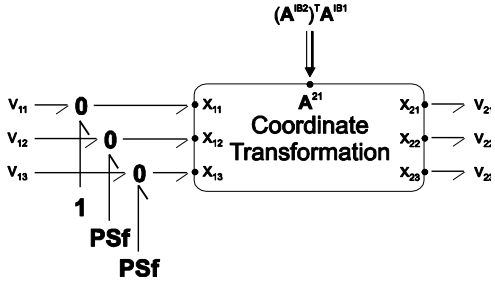


Figure A.15. Translational constraint YZ

APPENDIX B

PARAMETERS OF THE HMMWV MODEL

Table B.1. Parameters of the chassis

Description	Value
Moment of inertia	$\begin{bmatrix} 1504.55 & 0 & 0 \\ 0 & 5950.4 & 0 \\ 0 & 0 & 6357 \end{bmatrix}$ $\text{kg} \cdot \text{m}^2$
Mass	3514 kg
Front right upper A-arm connection point	$\begin{bmatrix} 2.032 & -0.4455 & -0.627 \end{bmatrix}^T$ m
Rear left upper A-arm connection point	$\begin{bmatrix} -1.27 & 0.4455 & -0.627 \end{bmatrix}^T$ m
Front left upper A-arm connection point	$\begin{bmatrix} 2.032 & 0.4455 & -0.627 \end{bmatrix}^T$ m
Front left lower A-arm connection point	$\begin{bmatrix} 2.032 & 0.2275 & -0.897 \end{bmatrix}^T$ m
Front right lower A-arm connection point	$\begin{bmatrix} 2.032 & -0.2275 & -0.897 \end{bmatrix}^T$ m
Rear left lower A-arm connection point	$\begin{bmatrix} -1.27 & 0.2275 & -0.897 \end{bmatrix}^T$ m
Rear right upper A-arm connection point	$\begin{bmatrix} -1.27 & -0.4455 & -0.627 \end{bmatrix}^T$ m
Rear right lower A-arm connection point	$\begin{bmatrix} -1.27 & -0.2275 & -0.897 \end{bmatrix}^T$ m
Idler arm connection point	$\begin{bmatrix} 1.647 & -0.28 & -0.762 \end{bmatrix}^T$ m
Pitman arm connection point	$\begin{bmatrix} 1.647 & 0.25 & -0.762 \end{bmatrix}^T$ m

Table B.2. Parameters of the front suspensions

Description	Value
Upper A-arm moment of inertia	$\begin{bmatrix} 0.0114 & 0 & 0 \\ 0 & 0.0104 & 0 \\ 0 & 0 & 0.0205 \end{bmatrix}$ kg·m^2
Upper A-arm mass	4.704 kg
Upper A-arm to chassis connection point	$\begin{bmatrix} 0 & -0.0778 & 0 \end{bmatrix}^T$ m
Upper A-arm to wheel hub connection point	$\begin{bmatrix} 0 & 0.1222 & -0.02 \end{bmatrix}^T$ m
Lower A-arm moment of inertia	$\begin{bmatrix} 0.1976 & 0 & 0 \\ 0 & 0.2158 & 0 \\ 0 & 0 & 0.1106 \end{bmatrix}$ kg·m^2
Lower A-arm mass	34.351 kg
Lower A-arm to wheel hub connection point	$\begin{bmatrix} 0 & 0.2846 & 0.02 \end{bmatrix}^T$ m
Lower A-arm to suspension strut connection point	$\begin{bmatrix} 0 & 0.0846 & 0 \end{bmatrix}^T$ m
Lower A-arm to chassis connection point	$\begin{bmatrix} 0 & -0.1704 & 0 \end{bmatrix}^T$ m
Wheel hub moment of inertia	$\begin{bmatrix} 0.1976 & 0 & 0 \\ 0 & 0.2158 & 0 \\ 0 & 0 & 0.1106 \end{bmatrix}$ kg·m^2
Wheel hub mass	34.351 kg
Wheel hub to upper A-arm connection point	$\begin{bmatrix} 0 & -0.0505 & 0.115 \end{bmatrix}^T$ m
Wheel hub to lower A-arm connection point	$\begin{bmatrix} 0 & -0.0135 & -0.115 \end{bmatrix}^T$ m
Wheel hub to tire connection point	$\begin{bmatrix} 0 & 0.0635 & 0 \end{bmatrix}^T$ m
Wheel hub to tie rod connection point	$\begin{bmatrix} -0.1354 & -0.0251 & 0 \end{bmatrix}^T$ m
Suspension stiffness	250 kN/m
Suspension damping	22460 N·s/m

Table B.3. Parameters of the rear suspensions

Description	Value
Upper A-arm moment of inertia	$\begin{bmatrix} 0.0114 & 0 & 0 \\ 0 & 0.0104 & 0 \\ 0 & 0 & 0.0205 \end{bmatrix}$ kg·m^2
Upper A-arm mass	4.704 kg
Upper A-arm to chassis connection point	$\begin{bmatrix} 0 & -0.0778 & 0 \end{bmatrix}^T$ m
Upper A-arm to wheel hub connection point	$\begin{bmatrix} 0 & 0.1222 & -0.02 \end{bmatrix}^T$ m
Lower A-arm moment of inertia	$\begin{bmatrix} 0.1976 & 0 & 0 \\ 0 & 0.2158 & 0 \\ 0 & 0 & 0.1106 \end{bmatrix}$ kg·m^2
Lower A-arm mass	34.351 kg
Lower A-arm to wheel hub connection point	$\begin{bmatrix} 0 & 0.2846 & 0.02 \end{bmatrix}^T$ m
Lower A-arm to suspension strut connection point	$\begin{bmatrix} 0 & 0.0846 & 0 \end{bmatrix}^T$ m
Lower A-arm to chassis connection point	$\begin{bmatrix} 0 & -0.1704 & 0 \end{bmatrix}^T$ m
Wheel hub moment of inertia	$\begin{bmatrix} 0.1976 & 0 & 0 \\ 0 & 0.2158 & 0 \\ 0 & 0 & 0.1106 \end{bmatrix}$ kg·m^2
Wheel hub mass	34.351 kg
Wheel hub to upper A-arm connection point	$\begin{bmatrix} 0 & -0.0505 & 0.115 \end{bmatrix}^T$ m
Wheel hub to lower A-arm connection point	$\begin{bmatrix} 0 & -0.0135 & -0.115 \end{bmatrix}^T$ m
Wheel hub to tire connection point	$\begin{bmatrix} 0 & 0.0635 & 0 \end{bmatrix}^T$ m
Suspension stiffness	300 kN/m
Suspension damping	35025 N·s/m

Table B.4. Parameters of the anti-roll bars

Description	Value
Front anti-roll bar stiffness	33.3333 kN/m
Rear anti-roll bar stiffness	66.6667 kN/m

Table B.5. Parameters of the steering mechanism

Description	Value
Idler/Pitman arm moment of inertia	$\begin{bmatrix} 0.0566 & 0 & 0 \\ 0 & 1.88 & 0 \\ 0 & 0 & 2.15 \end{bmatrix} \times 10^{-2}\ \mathrm{kg \cdot m^2}$
Idler/Pitman arm mass	2 kg
Idler/Pitman arm to steering link connection point	$\begin{bmatrix} 0.051 & 0 & 0 \end{bmatrix}^T$ m
Idler/Pitman arm to chassis connection point	$\begin{bmatrix} -0.051 & 0 & 0 \end{bmatrix}^T$ m
Steering link moment of inertia	$\begin{bmatrix} 0.204 & 0 & 0 \\ 0 & 8.33 \times 10^{-4} & 0 \\ 0 & 0 & 0.205 \end{bmatrix}\ \mathrm{kg \cdot m^2}$
Steering link mass	5 kg
Steering link to right tie rod connection point	$\begin{bmatrix} 0.0629 & -0.35 & 0 \end{bmatrix}^T$ m
Steering link to pitman arm connection point	$\begin{bmatrix} -0.011 & 0.25 & 0 \end{bmatrix}^T$ m
Steering link to idler arm connection point	$\begin{bmatrix} -0.011 & -0.28 & 0 \end{bmatrix}^T$ m
Steering link to left tie rod connection point	$\begin{bmatrix} 0.0629 & 0.35 & 0 \end{bmatrix}^T$ m
Right tie rod moment of inertia	$\begin{bmatrix} 1.85 & 0 & 0 \\ 0 & 0.08 & 0 \\ 0 & 0 & 1.85 \end{bmatrix} \times 10^{-2}\ \mathrm{kg \cdot m^2}$
Right tie rod mass	2 kg
Right tie rod to steering link connection point	$\begin{bmatrix} 0 & 0.1605 & 0 \end{bmatrix}^T$ m
Right tie rod to wheel hub connection point	$\begin{bmatrix} 0 & -0.1605 & 0 \end{bmatrix}^T$ m
Left tie rod moment of inertia	$\begin{bmatrix} 1.85 & 0 & 0 \\ 0 & 0.08 & 0 \\ 0 & 0 & 1.85 \end{bmatrix} \times 10^{-2}\ \mathrm{kg \cdot m^2}$
Left tie rod mass	2 kg
Left tie rod to steering link connection point	$\begin{bmatrix} 0 & -0.1605 & 0 \end{bmatrix}^T$ m
Left tie rod to wheel hub connection point	$\begin{bmatrix} 0 & 0.1605 & 0 \end{bmatrix}^T$ m

Table B.6. Parameters of the tires

Description	Value
Moment of inertia	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ kg·m^2
Mass	30 kg
Wheel hub connection point	$\begin{bmatrix} 0 & -0.15 & 0 \end{bmatrix}^T$ m
Damping	200 kN·s/m
Stiffness	1 MN/m
Radius	0.461 m
Cornering stiffness	100 kN
Friction coefficient	1.8

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] Ersal, T., Fathy, H. K., Louca, L. S., Rideout, D. G., and Stein, J. L., 2007, "A Review of Proper Modeling Techniques", IMECE2007-42031, *Proceedings of 2007 ASME International Mechanical Engineering Congress and Exposition*, American Society of Mechanical Engineers, New York, NY.

[2] Ersal, T., Fathy, H. K., and Stein, J. L., 2006, "Structural Simplification of Modular Bond-Graph Models Based on Junction Inactivity", IMECE2006-14567, *Proceedings of 2006 ASME International Mechanical Engineering Congress and Exposition*, American Society of Mechanical Engineers, New York, NY.

[3] Ersal, T., Fathy, H. K., and Stein, J. L., "Structural Simplification of Modular Bond-Graph Models Based on Junction Inactivity", Simulation Modelling Practice and Theory (in revision),(submitted to the Special Issue on Bond Graph Modelling).

[4] Ersal, T., Fathy, H. K., and Stein, J. L., 2006, "Orienting Body Reference Frames Using Karhunen-Loeve Expansion for More Effective Structural Simplification", IMECE2006-14572, *Proceedings of 2006 ASME International Mechanical Engineering Congress and Exposition*, ASME, New York, NY.

[5] Ersal, T., Fathy, H. K., and Stein, J. L., "Orienting Body Coordinate Frames Using Karhunen-Loève Expansion for More Effective Structural Simplification", Simulation Modelling Practice and Theory (in print),(submitted to the Special Issue on Bond Graph Modelling).

[6] Ersal, T., Fathy, H. K., and Stein, J. L., 2007, "Realization-Preserving Structure and Order Reduction of Nonlinear Energetic System Models Using Energy Trajectory Correlations", IMECE2007-42041, *Proceedings of 2007 ASME International Mechanical Engineering Congress and Exposition*, American Society of Mechanical Engineers, New York, NY.

[7] Ersal, T., Fathy, H. K., and Stein, J. L., "Realization-Preserving Structure and Order Reduction of Nonlinear Energetic System Models Using Energy Trajectory Correlations", Journal of Dynamic Systems Measurement and Control (in review).

[8] Ersal, T., 2003, "A Modular Modeling System for Reconfigurable Machine Tools", Master's Thesis, University of Michigan, Ann Arbor.

[9] Ersal, T., Stein, J. L., and Louca, L. S., 2004, "A Bond Graph Based Modular Modeling Approach Towards an Automated Modeling Environment for Reconfigurable Machine Tools", *Proceedings of IMAACA '04, Bond Graph Techniques for Modeling Dynamic Systems*, SCS.

[10] Ersal, T., Stein, J. L., and Louca, L. S., 2004, "A Modular Modeling Approach for the Design of Reconfigurable Machine Tools", IMECE2004-59806, *Proceedings of 2004 International Mechanical Engineering Congress and Exposition*, American Society of Mechanical Engineers, New York, NY.

[11] Oden, J. T., 2006, "Simulation-Based Engineering Science: Revolutionizing Engineering Science through Simulation", Report of the National Science Foundation Blue Ribbon Panel on Simulation-Based Engineering Science, National Science Foundation.

[12] Wilson, B. H. and Stein, J. L., 1995, "An Algorithm for Obtaining Proper Models of Distributed and Discrete Systems", Journal of Dynamic Systems, Measurement and Control, **117**(4), pp. 534-540.

[13] Bay, J. S., 1998, *Fundamentals of Linear State Space Systems*, McGraw-Hill.

[14] Kalman, R. E., 1963, "Mathematical Description of Linear Dynamical Systems", Society of Industrial and Applied Mathematics -- Journal on Control Series A, **1**(2), pp. 152-191.

[15] Wittenburg, J., 1977, *Dynamics of Systems of Rigid Bodies*, Teubner, Stuttgart.

[16] Schiehlen, W., 1991, "Computational Aspects in Multibody System Dynamics", *Proceedings of 2nd World Congress on Computational Mechanics*, **90**, pp. 569-582.

[17] Sueur, C. and Dauphin-Tanguy, G., 1991, "Bond Graph Approach to Multi-Time Scale Systems Analysis", Journal of the Franklin Institute, **328**(5-6), pp. 1005-1026.

[18] Orbak, A. Y., Turkay, O. S., Eskinat, E., and Youcef-Toumi, K., 2003, "Model Reduction in the Physical Domain", Proceedings of the Institution of Mechanical Engineers. Part I: Journal of Systems and Control Engineering, **217**(6), pp. 481-496.

[19] Rideout, D. G., Stein, J. L., and Louca, L. S., 2004, "System Partitioning and Physical- Domain Model Reduction through Assessment of Bond Graph Junction Structure", *Proceedings of IMAACA '04, Bond Graph Techniques for Modeling Dynamic Systems*, SCS.

[20] Louca, L. S., Stein, J. L., Hulbert, G. M., and Sprague, J., 1997, "Proper Model Generation: An Energy-Based Methodology", *Proceedings of 1997 International Conference on Bond Graph Modeling*, SCS, **29**, pp. 44-49.

[21] Karnopp, D. and Rosenberg, R. C., 1975, *System Dynamics: A Unified Approach*, John Wiley & Sons.

[22] Elrazaz, Z. and Sinha, N. K., 1981, "A Review of Some Model Reduction Techniques", Canadian Electrical Engineering Journal, **6**(1), pp. 34-40.

[23] Lamba, S. S. and Mahmoud, M. S., 1982, "Model Simplification - an Overview", *Proceedings of Theory and Application of Digital Control, Proceedings of the IFAC Symposium.*, IFAC by Pergamon Press, Oxford, Engl, pp. 479-487.

[24] Saksena, V. R., O'Reilly, J., and Kokotovic, P. V., 1984, "Singular Perturbations and Time-Scale Methods in Control Theory: Survey 1976-83", Automatica, **20**(3), pp. 273-293.

[25] Bertrand, P., Duc, G., and Michailesco, G., 1985, "Développements Récents Sur La Réduction de Modèles", Automatique-Productique Informatique Industrielle, **19**(2), pp. 131-146.

[26] Sugimoto, S., Kaino, E., and Mori, Y., 1985, "Comparative Studies for Several Model Reduction Algorithms", *Proceedings of International Conference - Control 85 (IEE Conf. Publ. No. 252)*, IEE, pp. 673-678.

[27] Fortuna, L., Nunnari, G., and Gallo, A., 1992, *Model Order Reduction Techniques with Applications in Electrical Engineering*, Springer-Verlag, London.

[28] Gugercin, S. and Antoulas, A. C., 2000, "A Comparative Study of 7 Algorithms for Model Reduction", *Proceedings of 39th IEEE Conference on Decision and Control*, Institute of Electrical and Electronics Engineers Inc., **3**, pp. 2367-2372.

[29] Antoulas, A. C., Sorensen, D. C., and Gugercin, S., 2001, "A Survey of Model Reduction Methods for Large Scale Systems", Contemporary Mathematics, **280**, pp. 193-219.

[30] Antoulas, A. C. and Gugercin, S., 2002, "A New Approach to Model Reduction Which Preserves Stability and Passivity: An Overview", *Proceedings of IEEE Conference on Decision and Control*, Institute of Electrical and Electronics Engineers Inc., **3**, pp. 2544-2545.

[31] Gugercin, S. and Antoulas, A. C., 2004, "A Survey of Model Reduction by Balanced Truncation and Some New Results", International Journal of Control, **77**(8), pp. 748-766.

[32] Qu, Z.-Q., 2004, *Model Order Reduction Techniques : With Applications in Finite Element Analysis*, Springer, London ; New York.

[33] Davison, E. J., 1966, "A Method for Simplifying Linear Dynamic Systems", IEEE Transactions on Automatic Control, **AC-11**(1), pp. 93-101.

[34] Marshall, S. A., 1966, "An Approximate Method for Reducing the Order of a Linear System", Control, **10**, pp. 642-643.

[35] Chidambara, M. R., 1967, "On "A Method for Simplifying Linear Dynamic Systems"", IEEE Transactions on Automatic Control, **AC-12**(6), pp. 119-121.

[36] Chidambara, M. R., 1967, "Further Comments On "A Method for Simplifying Linear Dynamic Systems"", IEEE Transactions on Automatic Control, **AC-12**(6), pp. 799-800.

[37] Chidambara, M. R. and Davison, E. J., 1967, "Further Remarks on Simplifying Linear Dynamic Systems", IEEE Transactions on Automatic Control, **AC-12**(2), pp. 213-214.

[38] Davison, E. J., 1968, "A New Method for Simplifying Large Linear Dynamic Systems", IEEE Transactions on Automatic Control, **AC-13**(2), pp. 215-216.

[39] Fossard, A., 1970, "On a Method for Simplifying Linear Dynamic Systems", IEEE Transactions on Automatic Control, **AC-15**(2), pp. 261-262.

[40] Nicholson, H., 1964, "Dynamic Optimisation of a Boiler", Proceedings of the Institution of Electrical Engineers, **1ll**(8), pp. 1479-1499.

[41] Mitra, D., 1967, "On the Reduction of Complexity of Linear Dynamic Models", Technical Report, AEEW-R520, United Kingdom Atomic Energy Authority, Winfrith.

[42] Mitra, D., 1969, "The Reduction of Complexity of Linear Time Invariant Dynamical Systems", *Proceedings of IFAC World Congress*, pp. 19-33.

[43] Aoki, M., 1968, "Control of Large-Scale Dynamic Systems by Aggregation", IEEE Transactions on Automatic Control, **AC-13**(3), pp. 246-253.

[44] Hickin, J. and Sinha, N. K., 1975, "Aggregation Matrices for a Class of Low-Order Models for Large-Scale Systems", Electronics Letters, **11**(9), pp. 186.

[45] Siret, J. M., Michailesco, G., and Bertrand, P., 1977, "Representation of Linear Dynamical Systems by Aggregated Models", International Journal of Control, **26**(1), pp. 121-128.

[46] Hickin, J., 1978, "A Unified Theory of Model Reduction for Linear Time Invariant Dynamical Systems", Ph.D. Dissertation, McMaster University, Hamilton, Ontario, Canada.

[47] Aoki, M., 1978, "Some Approximation Methods for Estimation and Control of Large Scale Systems", IEEE Transactions on Automatic Control, **AC-23**(2), pp. 173-182.

[48] Hickin, J. and Sinha, N. K., 1980, "Model Reduction for Linear Multivariable Systems", IEEE Transactions on Automatic Control, **AC-25**(6), pp. 1121-1127.

[49] Prandtl, L., 1904, "Über Flüssigkeitsbewegung Bei Sehr Kleiner Reibung", *Proceedings of Verhandlung des III. Internationalen Mathematiker Kongresses*.

[50] Tikhonov, A., 1948, "On the Dependence of the Solutions of Differential Equations on a Small Parameter", Mat. Sb., **22**, pp. 193-204.

[51] Levinson, N., 1950, "Perturbations of Discontinuous Solutions of Non-Linear Systems of Differential Equations", Acta Mathematica, **82**, pp. 71-106.

[52] Vasileva, A. B., 1963, "Asymptotic Behavior of Solutions to Certain Problems Involving Nonlinear Differential Equations Containing a Small Parameter Multiplying the Highest Derivatives", Russian Mathematical Surveys, **18**, pp. 13-81.

[53] Wasow, W., 1965, *Asymptotic Expansions for Ordinary Differential Equations*, Wiley-Interscience, New York.

[54] Kokotovic, P. and Sannuti, P., 1968, "Singular Perturbation Method for Reducing Model Order in Optimal Control Design", IEEE Transactions on Automatic Control, **AC-13**(4), pp. 377-384.

[55] Kokotovic, P. V., O'Malley, R. E., Jr., and Sannuti, P., 1976, "Singular Perturbations and Order Reduction in Control Theory-an Overview", Automatica, **12**(2), pp. 123-132.

[56] Kokotovic, P. V., 1984, "Applications of Singular Perturbation Techniques to Control-Problems", Siam Review, **26**(4), pp. 501-550.

[57] Kokotovic, P., Khalil, H. K., and O'Reilly, J., 1986, *Singular Perturbation Methods in Control: Analysis and Design*, Academic Press, London.

[58] Liu, Y. and Anderson, B. D. O., 1989, "Singular Perturbation Approximation of Balanced Systems", International Journal of Control, **50**(4), pp. 1379-1405.

[59] Wilson, B. H. and Stein, J. L., 1992, "An Algorithm for Obtaining Minimum-Order Models of Distributed and Discrete Systems", *Proceedings of Winter Annual Meeting of the American Society of Mechanical Engineering*, ASME, New York, NY, USA, **41**, pp. 47-58.

[60] Ferris, J. B., Stein, J. L., and Bernitsas, M. M., 1994, "Development of Proper Models of Hybrid Systems", *Proceedings of ASME Winter Annual Meeting, Symposium on Automated Modeling*, pp. 629-636.

[61] Walker, D. G., Stein, J. L., and Ulsoy, A. G., 2000, "An Input-Output Criterion for Linear Model Deduction", Journal of Dynamic Systems, Measurement, and Control, **122**(3), pp. 507-513.

[62] Wilson, B. H. and Taylor, J. H., 1998, "A Frequency Domain Model-Order-Deduction Algorithm for Linear Systems", Transactions of the ASME. Journal of Dynamic Systems, Measurement and Control, **120**(2), pp. 185-192.

[63] Taylor, J. H. and Wilson, B. H., 1995, "A Frequency-Domain Model-Order-Deduction Algorithm for Nonlinear Systems", *Proceedings of International Conference on Control Applications*, IEEE, pp. 1053-1058.

[64] Pirvu, A.-M., Dauphin-Tanguy, G., and Kubiak, P., 2007, "Automatic System for Bond Graph Model Adaptation - Application to an Electro-Hydrostatic Actuator", *Proceedings of 2007 International Conference on Bond Graph Modeling and Simulation (ICBGM '07)*, J. Granda and F. Cellier, eds., SCS, pp. 35-41.

[65] Kisung, S., Zhun, F., Jianjun, H., Goodman, E. D., and Rosenberg, R. C., 2003, "Toward a Unified and Automated Design Methodology for Multi-Domain Dynamic Systems Using Bond Graphs and Genetic Programming", Mechatronics, **13**(8-9), pp. 851-885.

[66] Fan, Z., Seo, K., Hu, J., Goodman, E. D., and Rosenberg, R. C., 2004, "A Novel Evolutionary Engineering Design Approach for Mixed-Domain Systems", Engineering Optimization, **36**(2), pp. 127-147.

[67] Margolis, D. L. and Young, G. E., 1977, "Reduction of Models of Large Scale Lumped Structures Using Normal Modes and Bond Graphs", Journal of the Franklin Institute, **304**(1), pp. 65-79.

[68] Fertis, D. G., 1995, *Mechanical and Structural Vibrations*, John Wiley & Sohns, New York.

[69] Skogestad, S. and Postlethwaite, I., 1996, *Multivariable Feedback Control: Analysis and Design*, John Wiley and Sons.

[70] Mohanty, P. and Rixen, D. J., 2004, "Operational Modal Analysis in the Presence of Harmonic Excitation", Journal of Sound and Vibration, **270**(1-2), pp. 93-109.

[71] Vukazich, S. M., Mish, K. D., and Romstad, K. M., 1996, "Nonlinear Dynamic Response of Frames Using Lanczos Modal Analysis", Journal of Structural Engineering, **122**(12), pp. 1418-1426.

[72] Li, A. and Dowell, E. H., 2004, "Asymptotic Modal Analysis of Dynamical Systems: The Effect of Modal Cross-Correlation", Journal of Sound and Vibration, **276**(1-2), pp. 65-80.

[73] Hurty, W. C., 1960, "Vibrations of Structural Systems by Component Mode Synthesis", American Society of Civil Engineers -- Proceedings-- Journal of the Engineering Mechanics Division, **86**(EM4, Part 1), pp. 51-69.

[74] Hurty, W. C., 1965, "Dynamic Analysis of Structural Systems Using Component Modes", AIAA Journal, **3**(4), pp. 678-685.

[75] Craig, J., R.R. and Bampton, M. C. C., 1968, "Coupling of Substructures for Dynamic Analyses", AIAA Journal, **6**(7), pp. 1313-1319.

[76] Aoyama, Y. and Yagawa, G., 2001, "Large-Scale Eigenvalue Analysis of Structures Using Component Mode Synthesis", JSME International Journal, Series A: Solid Mechanics and Material Engineering, **44**(4), pp. 631-640.

[77] Apiwattanalunggarn, P., Shaw, S. W., and Pierre, C., 2005, "Component Mode Synthesis Using Nonlinear Normal Modes", Nonlinear Dynamics, **41**(1-3), pp. 17-46.

[78] Moon, J. D. and Cho, D. W., 1992, "A Component Mode Synthesis Applied to Mechanisms for an Investigation of Vibration", Journal of Sound and Vibration, **157**(1), pp. 67-79.

[79] Sung, S. H. and Nefske, D. J., 1986, "Component Mode Synthesis of a Vehicle Structural-Acoustic System Model", AIAA Journal, **24**(6), pp. 1021-1026.

[80] Verros, G. and Natsiavas, S., 2002, "Ride Dynamics of Nonlinear Vehicle Models Using Component Mode Synthesis", Journal of Vibration and Acoustics, Transactions of the ASME, **124**(3), pp. 427-434.

[81] Karpel, M., Moulin, B., and Feldgun, V., 2003, "Component Mode Synthesis of a Vehicle System Model Using the Fictitious Mass Method", *Proceedings of 44th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials*

Conference, Apr 7-10 2003, American Inst. Aeronautics and Astronautics Inc., **3**, pp. 1836-1845.

[82] Arnoldi, W. E., 1951, "The Principle of Minimized Iterations in the Solution of the Matrix Eigenvalue Problem", Quarterly of Applied Mathematics, **9**(1), pp. 17-29.

[83] Sorensen, D. C., 1992, "Implicit Application of Polynomial Filters in a K-Step Arnoldi Method", Siam Journal on Matrix Analysis and Applications, **13**(1), pp. 357-385.

[84] Lanczos, C., 1950, "An Iteration Method for the Solution of the Eigenvalue Problem of Linear Differential and Integral Operators", Journal of Research of the National Bureau of Standards, **45**(4), pp. 255-282.

[85] Gallivan, K., Grimme, E., and Van Dooren, P., 1994, "Pade Approximation of Large-Scale Dynamic Systems with Lanczos Methods", *Proceedings of 33rd IEEE Conference on Decision and Control. Part 1 (of 4)*, IEEE, Piscataway, NJ, USA, **1**, pp. 443-448.

[86] Ruhe, A., 1994, "Rational Krylov Algorithms for Nonsymmetric Eigenvalue Problems II: Matrix Pairs", Linear Algebra and Its Applications, **197**, pp. 283-295.

[87] Grimme, E. J., 1997, "Krylov Projection Methods for Model Reduction", Ph.D. Dissertation, University of Illinois, Urbana-Champaign.

[88] Alexandro, F. J., Jr., 1984, "Stable Partial Padé Approximations for Reduced-Order Transfer Functions", IEEE Transactions on Automatic Control, **AC-29**(2), pp. 159-162.

[89] Goldman, M. J., Porras, W. J., and Leondes, C. T., 1981, "Multivariable Systems Reduction Via Cauer Forms", International Journal of Control, **34**(4), pp. 623-650.

[90] Shieh, L.-S., Chang, F.-R., and Yates, R. E., 1986, "The Generalized Matrix Continued-Fraction Descriptions and Their Application to Model Simplification", International Journal of Systems Science, **17**(1), pp. 1-19.

[91] Chen, C. F., 1974, "Model Reduction of Multivariable Control Systems by Means of Matrix Continued Fractions", International Journal of Control, **20**(2), pp. 225-238.

[92] Chen, C. F. and Shieh, L. S., 1968, "A Novel Approach to Linear Model Simplification", International Journal of Control, **8**(6), pp. 561-570.

[93] Chen, C. F. and Shieh, L. S., 1969, "Continued Fraction Inversion by Routh's Algorithm", IEEE Transactions on Circuit Theory, **CT-16**(2), pp. 197-202.

[94] Hutton, M. F. and Friedland, B., 1975, "Routh Approximations for Reducing Order of Linear, Time-Invariant Systems", IEEE Transactions on Automatic Control, **AC-20**(3), pp. 329-337.

[95] Chen, T. C., Han, K. W., and Chang, C. Y., 1979, "Reduction of Transfer Functions by the Stability-Equation Method", Journal of the Franklin Institute, **308**(4), pp. 389-404.

[96] Shamash, Y., 1975, "Model Reduction Using the Routh Stability Criterion and the Pade Approximation Technique", International Journal of Control, **21**(3), pp. 475-484.

[97] Pal, J., 1979, "Stable Reduced-Order Padé Approximation Using the Routh-Hurwitz Array", Electronics Letters, **15**(8), pp. 225-226.

[98] Pal, J. and Ray, L. M., 1980, "Stable Padé Approximants to Multivariable Systems Using a Mixed Method", Proceedings of the IEEE, **68**(1), pp. 176-178.

[99] Chen, T. C., Chang, C. Y., and Han, K. W., 1980, "Model Reduction Using the Stability-Equation Method and the Padé Approximation Method", Journal of the Franklin Institute, **309**(6), pp. 473-490.

[100] Chen, T. C., Chang, C. Y., and Han, K. W., 1980, "Model Reduction Using the Stability-Equation Method and the Continued-Fraction Method", International Journal of Control, **32**(1), pp. 81-94.

[101] de Villemagne, C. and Skelton, R. E., 1987, "Model Reductions Using a Projection Formulation", International Journal of Control, **46**(6), pp. 2141-2169.

[102] Adamjan, V. M., Arov, D. V., and Krein, M. G., 1971, "Analytic Properties of Schmidt Pairs for a Hankel Operator and the Generalized Schmidt-Takagi Problem", Math. USSR Sbornik, **15**(1).

[103] Kung, S., 1978, "A New Identification and Model Reduction Algorithm Via Singular Value Decompositions", *Proceedings of Twelfth Asilomar Conference on Circuits, Systems and Computers*, IEEE, pp. 705-714.

[104] Kung, S.-Y. and Lin, D. W., 1981, "Optimal Hankel-Norm Model Reductions: Multivariable Systems", IEEE Transactions on Automatic Control, **AC-26**(4), pp. 832-852.

[105] Glover, K., 1984, "All Optimal Hankel-Norm Approximations of Linear Multivariable Systems and Their L-Infinity Error Bounds", International Journal of Control, **39**(6), pp. 1115-1193.

[106] Safonov, M. G., Chiang, R. Y., and Limebeer, D. J. N., 1990, "Optimal Hankel Model Reduction for Nonminimal Systems", IEEE Transactions on Automatic Control, **35**(4), pp. 496-502.

[107] Ionescu, V. and Oara, C., 2001, "The Four-Block Adamjan-Arov-Krein Problem for Discrete-Time Systems", Linear Algebra and Its Applications, **328**(1-3), pp. 95-119.

[108] Gu, G., 2005, "All Optimal Hankel-Norm Approximations and Their L-Infinity Error Bounds in Discrete-Time", International Journal of Control, **78**(6), pp. 408-423.

[109] Ball, J. A. and Ran, A. C. M., 1987, "Optimal Hankel Norm Model Reductions and Wiener-Hopf Factorization. I: The Canonical Case", SIAM Journal on Control and Optimization, **25**(2), pp. 362-382.

[110] Ionescu, V. and Oara, C., 1996, "Class of Suboptimal and Optimal Solutions to the Discrete Nehari Problem: Characterization and Computation", International Journal of Control, **64**(3), pp. 483-509.

[111] Kung, S.-Y. and Lin, D. W., 1981, "A State-Space Formulation for Optimal Hankel-Norm Approximations", IEEE Transactions on Automatic Control, **AC-26**(4), pp. 942-6.

[112] Benner, P., Quintana-Orti, E. S., and Quintana-Orti, G., 2004, "Computing Optimal Hankel Norm Approximations of Large-Scale Systems", *Proceedings of 43rd IEEE Conference on Decision and Control (CDC)*, Institute of Electrical and Electronics Engineers Inc., Piscataway, NJ 08855-1331, United States, **3**, pp. 3078-3083.

[113] Loève, M. M., 1955, *Probability Theory*, Princeton, N.J., VanNostrand.

[114] Karhunen, K., 1946, "Zur Spektraltheorie Stochastischer Prozesse", Annales Academiae Scientarum Fennicae, **37**.

[115] Hotelling, H., 1933, "Analysis of a Complex of Statistical Variables into Principal Components", Journal of Educational Psychology, **24**, pp. 417-441, 498-520.

[116] Lorenz, E. N., 1956, "Empirical Orthogonal Functions and Statistical Weather Prediction", Scientific Report, 1, MIT, Department of Meteorology, Cambridge.

[117] Lumley, J. L., "The Structure of Inhomogeneous Turbulent Flow", in *Atmospheric Turbulence and Radio Wave Propagation*, A. M. Yaglom and V. I. Tatarski, Eds. Nauka, Moscow, 1967, pp. 166-178.

[118] Golub, G. H. and Van Loan, C. F., 1983, *Matrix Computations*, North Oxford Academic, Oxford.

[119] Sirovich, L., 1987, "Turbulence and the Dynamics of Coherent Structures. I. Coherent Structures", Quarterly of Applied Mathematics, **45**(3), pp. 561-570.

[120] Sirovich, L., 1987, "Turbulence and the Dynamics of Coherent Structures. II. Symmetries and Transformations", Quarterly of Applied Mathematics, **45**(3), pp. 573-582.

[121] Sirovich, L., 1987, "Turbulence and the Dynamics of Coherent Structures. III. Dynamics and Scaling", Quarterly of Applied Mathematics, **45**(3), pp. 583-590.

[122] Brooks, C. L., Karplus, M., and Pettitt, B. M., 1988, *Proteins: A Theoretical Perspective of Dynamics, Structure and Thermodynamics*, Wiley, New York.

[123] Berkooz, G., Holmes, P., and Lumley, J. L., 1993, "Proper Orthogonal, Decomposition in the Analysis of Turbulent Flows", Annual Review of Fluid Mechanics, **25**, pp. 537.

[124] Kosambi, D. D., 1943, "Statistics in Function Space", Journal of the Indian Mathematical Society, **7**, pp. 76-88.

[125] Loève, M. M., 1945, "Functions Aléatoire de Second Ordre", Comptes Rendus de l'Académie des Sciences, pp. 220.

[126] Pougachev, V. S., 1953, "General Theory of the Correlations of Random Functions", Izvestiya Rossiiskoi Akademii Nauk, Seriya Matematicheskaya, **17**, pp. 1401-1402.

[127] Obukhov, A. M., 1954, "Statistical Description of Continuous Fields", Trudy Geofizicheskogo Instituta, Akademiya Nauk SSSR, **24**, pp. 3-42.

[128] Holmes, P., Lumley, J. L., and Berkooz, G., 1996, *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*, Cambridge University Press, Cambridge.

[129] Papoulis, A., 1965, *Probablity, Random Variables, and Stochastic Processes*, McGraw-Hill, New York.

[130] Rosenfeld, A. and Kak, A. C., 1982, *Digital Picture Processing*, Academic, New York.

[131] Algazi, V. R. and Sakrison, D. J., 1969, "On the Optimality of the Karhunen-Loeve Expansion", IEEE Transactions on Information Theory, **IT-15**(2), pp. 319-321.

[132] Andrews, C. A., Davies, J. M., and Schwarz, G. R., 1967, "Adaptive Data Compression", IEEE Proceedings, **55**(3), pp. 267-277.

[133] Preisendorfer, R. W., 1988, *Principal Component Analysis in Meteorology and Oceanography*, Elsevier, Amsterdam.

[134] Gay, D. H. and Ray, W. H., 1995, "Identification and Control of Distributed Parameter Systems by Means of the Singular Value Decomposition", Chemical Engineering Science, **50**(10), pp. 1519-1539.

[135] Wu, G. G., Liang, Y. C., Lin, W. Z., Lee, H. P., and Lim, S. P., 2003, "A Note on Equivalence of Proper Orthogonal Decomposition Methods", Journal of Sound and Vibration, **265**(5), pp. 1103-1110.

[136] Utku, S., Clemente, J. L. M., and Salama, M., 1985, "Errors in Reduction Methods", Computers and Structures, **21**(6), pp. 1153-1157.

[137] Kunisch, K. and Volkwein, S., 2002, "Galerkin Proper Orthogonal Decomposition Methods for a General Equation in Fluid Dynamics", SIAM Journal on Numerical Analysis, **40**(2), pp. 492-515.

[138] Rathinam, M. and Petzold, L. R., 2003, "A New Look at Proper Orthogonal Decomposition", SIAM Journal on Numerical Analysis, **41**(5), pp. 1893-1925.

[139] Homescu, C., Petzold, L. R., and Serban, R., 2005, "Error Estimation for Reduced-Order Models of Dynamical Systems", SIAM Journal on Numerical Analysis, **43**(4), pp. 1693-1714.

[140] Homescu, C., Petzold, L. R., and Serban, R., 2007, "Error Estimation for Reduced-Order Models of Dynamical Systems", SIAM Review, **49**(2), pp. 277-299.

[141] Moore, B. C., 1981, "Principal Component Analysis in Linear Systems: Controllability, Observability, and Model Reduction", IEEE Transactions on Automatic Control, **26**(1), pp. 17-32.

[142] Kalman, R. E., 1965, "Irreducible Realizations and Degree of a Rational Matrix", Journal of the Society for Industrial and Applied Mathematics, **13**(2), pp. 520-544.

[143] Kabamba, P. T., 1985, "Balanced Gains and Their Significance for L2 Model Reduction", IEEE Transactions on Automatic Control, **AC-30**(7), pp. 690-693.

[144] Penzl, T., 1999, "Cyclic Low-Rank Smith Method for Large Sparse Lyapunov Equations", SIAM Journal of Scientific Computing, **21**(4), pp. 1401-1418.

[145] Sorensen, D. C. and Antoulas, A. C., 2002, "The Sylvester Equation and Approximate Balanced Reduction", Linear Algebra and Its Applications, **351-352**, pp. 671-700.

[146] Gugercin, S., Sorensen, D. C., and Antoulas, A. C., 2003, "A Modified Low-Rank Smith Method for Large-Scale Lyapunov Equations", Numerical Algorithms, **32**(1), pp. 27-55.

[147] Desai, U. B. and Pal, D., 1984, "A Transformation Approach to Stochastic Model Reduction", IEEE Transactions on Automatic Control, **AC-29**(12), pp. 1097-1100.

[148] Green, M., 1988, "A Relative Error Bound for Balanced Stochastic Truncation", IEEE Transactions on Automatic Control, **33**(10), pp. 961-965.

[149] Green, M., 1988, "Balanced Stochastic Realizations", Linear Algebra and Its Applications, **98**, pp. 211-247.

[150] Opdenacker, P. C. and Jonckheere, E. A., 1988, "A Contraction Mapping Preserving Balanced Reduction Scheme and Its Infinity Norm Error Bounds", IEEE Transactions on Circuits and Systems, **35**(2), pp. 184-189.

[151] Jonckheere, E. A. and Silverman, L. M., 1983, "A New Set of Invariants for Linear Systems - Application to Reduced Order Compensator Design", IEEE Transactions on Automatic Control, **AC-28**(10), pp. 953-964.

[152] Enns, D. F., 1984, "Model Reduction with Balanced Realizations: An Error Bound and a Frequency Weighted Generalization", *Proceedings of 23rd IEEE Conference on Decision and Control*, IEEE, New York, NY, USA, pp. 127-132.

[153] Lin, C.-A. and Chiu, T.-Y., 1992, "Model Reduction Via Frequency Weighted Balanced Realization", Control Theory and Advanced Technology, **8**(2), pp. 341-351.

[154] Sreeram, V. and Anderson, B. D. O., 1995, "Frequency Weighted Balanced Reduction Technique: A Generalization and an Error Bound", *Proceedings of 34th IEEE Conference on Decision and Control. Part 4 (of 4), Dec 13-15 1995*, IEEE, Piscataway, NJ, USA, **4**, pp. 3576-3581.

[155] Wang, G., Sreeram, V., and Liu, W. Q., 1999, "New Frequency-Weighted Balanced Truncation Method and an Error Bound", IEEE Transactions on Automatic Control, **44**(9), pp. 1734-1737.

[156] Zhou, K., 1993, "Frequency-Weighted Model Reduction with L Infinity Error Bound", Systems & Control Letters, **21**(2), pp. 115-125.

[157] Hahn, J. and Edgar, T. F., 2002, "Balancing Approach to Minimal Realization and Model Reduction of Stable Nonlinear Systems", Industrial and Engineering Chemistry Research, **41**(9), pp. 2204-2212.

[158] Scherpen, J. M. A. and Gray, W. S., 2000, "Minimality and Local State Decompositions of a Nonlinear State Space Realization Using Energy Functions", IEEE Transactions on Automatic Control, **45**(11), pp. 2079-2086.

[159] Scherpen, J. M. A., 1993, "Balancing for Nonlinear Systems", Systems & Control Letters, **21**(2), pp. 143-153.

[160] Lee, K. S., Eom, Y., Chung, J. W., Choi, J., and Yang, D., 2000, "A Control-Relevant Model Reduction Technique for Nonlinear Systems", Computers & Chemical Engineering, **24**(2-7), pp. 309-315.

[161] Hahn, J. and Edgar, T. F., 2002, "An Improved Method for Nonlinear Model Reduction Using Balancing of Empirical Gramians", Computers and Chemical Engineering, **26**(10), pp. 1379-1397.

[162] Lall, S., Marsden, J. E., and Glavaski, S., 2002, "A Subspace Approach to Balanced Truncation for Model Reduction of Nonlinear Control Systems", International Journal of Robust and Nonlinear Control, **12**(6), pp. 519-535.

[163] Liu, Z. and Wagner, J., 2002, "Nonlinear Model Reduction for Dynamic and Automotive System Descriptions", Journal of Dynamic Systems, Measurement and Control, Transactions of the ASME, **124**(4), pp. 637-647.

[164] Verriest, E. I. and Gray, W. S., 2001, "Nonlinear Balanced Realizations", *Proceedings of 40th IEEE Conference on Decision and Control*, Institute of Electrical and Electronics Engineers Inc., **4**, pp. 3250-3251.

[165] Ma, X. and De Abreu-Garcia, J. A., 1988, "On the Computation of Reduced Order Models of Nonlinear Systems Using Balancing Technique", *Proceedings of 27th IEEE Conference on Decision and Control*, IEEE, pp. 1165-1166.

[166] Salimbahrami, B. and Lohmann, B., 2004, "A Simulation-Free Nonlinear Model Order-Reduction Approach and Comparison Study", Mathematical and Computer Modelling of Dynamical Systems, **10**(3-4), pp. 317-329.

[167] Yousefi, A. and Lohmann, B., 2004, "Balancing and Optimization for Order Reduction of Nonlinear Systems", *Proceedings of American Control Conference*,

Institute of Electrical and Electronics Engineers Inc., Piscataway, NJ 08855-1331, United States, **1**, pp. 108-112.

[168] Fujimoto, K. and Tsubakino, D., 2006, "On Computation of Nonlinear Balanced Realization and Model Reduction", *Proceedings of 2006 American Control Conference*, IEEE, pp. 460-465.

[169] Skelton, R. E., 1980, "Cost Decomposition of Linear Systems with Application to Model Reduction", International Journal of Control, **32**(6), pp. 1031-1055.

[170] Skelton, R. E. and Yousuff, A., 1983, "Component Cost Analysis of Large Scale Systems", International Journal of Control, **37**(2), pp. 285-304.

[171] Skelton, R. E. and Kabamba, P., 1986, "Comments on 'Balanced Gains and Their Significance for L2 Model Reduction' by P. T. Kabamba", IEEE Transactions on Automatic Control, **AC-31**(8), pp. 796-797.

[172] Skelton, R. E., 1988, *Dynamic Systems Control: Linear Systems Analysis and Synthesis*, Wiley, New York.

[173] Skelton, R. E., 1989, "The Explicit Relation between Component Cost Analysis and Balanced Coordinates", *Proceedings of 28th IEEE Conference on Decision and Control*, IEEE, Piscataway, NJ, USA, **2**, pp. 1326-1330.

[174] Hyland, D. C. and Bernstein, D. S., 1985, "The Optimal Projection Equations for Model Reduction and the Relationships among the Methods of Wilson, Skelton, and Moore", IEEE Transactions on Automatic Control, **AC-30**(12), pp. 1201-1211.

[175] Wilson, D. A., 1970, "Optimum Solution of Model-Reduction Problem", *Proceedings of Institution of Electrical Engineers*, **117**, pp. 1161-1165.

[176] Wilson, D. A., 1974, "Model Reduction for Multivariable Systems", International Journal of Control, **20**(1), pp. 57-64.

[177] Galiana, F. D., 1973, "On the Approximation of Multiple Input-Multiple Output Constant Linear Systems", International Journal of Control, **17**(6), pp. 1313-1324.

[178] Riggs, J. B. and Edgar, T. F., 1974, "Least Squares Reduction of Linear Systems Using Impulse Response", International Journal of Control, **20**(2), pp. 213-223.

[179] Aplevich, J. D., 1973, "Approximation of Discrete Linear Systems", International Journal of Control, **17**(3), pp. 565-575.

[180] Hirzinger, G. and Kreisselmeier, G., 1975, "On Optimal Approximation of High-Order Linear Systems by Low-Order Models", International Journal of Control, **22**(3), pp. 339-408.

[181] Wilson, D. A. and Mishra, R. N., 1979, "Optimal Reduction of Multivariable Systems", International Journal of Control, **29**(2), pp. 267-278.

[182] Mishra, R. N. and Wilson, D. A., 1980, "A New Algorithm for Optimal Reduction of Multivariable Systems", International Journal of Control, **31**(3), pp. 443-466.

[183] Howitt, G. D. and Luus, R., 1990, "Model Reduction by Minimization of Integral Square Error Performance Indexes", Journal of the Franklin Institute-Engineering and Applied Mathematics, **327**(3), pp. 343-357.

[184] Krajewski, W., Lepschy, A., and Viaro, U., 1996, "An Optimization Approach to Model Reduction", *Proceedings of Third International Symposium on Methods and Models in Automation and Robotics*, Tech. Univ. Szczecin, **1**, pp. 351-356.

[185] Zhang, L. and Lam, J., 1999, "A Parametric Optimization Approach to Discrete-Time Model Reduction", *Proceedings of 14th World Congress of IFAC*, Elsevier Sci, **5**, pp. 105-110.

[186] Krajewski, W., Lepschy, A., Redivo-Zaglia, M., and Viaro, U., 1995, "A Program for Solving the L2 Reduced-Order Model Problem with Fixed Denominator Degree", Numerical Algorithms, **9**(3-4), pp. 355-377.

[187] Luus, R., 1980, "Optimization in Model Reduction", International Journal of Control, **32**(5), pp. 741-747.

[188] Sou, K. C., Megretski, A., and Daniel, L., 2005, "A Quasi-Convex Optimization Approach to Parameterized Model Order Reduction", *Proceedings of 42nd Design Automation Conference*, Institute of Electrical and Electronics Engineers Inc., Piscataway, NJ 08855-1331, United States, pp. 933-938.

[189] Gouda, M. M., Danaher, S., and Underwood, C. P., 2002, "Building Thermal Model Reduction Using Nonlinear Constrained Optimization", Building and Environment, **37**(12), pp. 1255-1265.

[190] Hachtel, G., Sangiovanni-Vincentelli, A., and Visvanathan, V., 1981, "An Optimization-Based Approach to Model Simplification", *Proceedings of IEEE International Symposium on Circuits and Systems*, IEEE, Piscataway, NJ, **3**, pp. 995-1000.

[191] Assunção, E. and Peres, P. L. D., 1999, "A Global Optimization Approach for the H2-Norm Model Reduction Problem", *Proceedings of 38th IEEE Conference on Decision and Control*, IEEE, Piscataway, NJ, USA, **2**, pp. 1857-1862.

[192] Chen, H.-F. and Fang, H.-T., 2002, "Nonconvex Stochastic Optimization for Model Reduction", Journal of Global Optimization, **23**(3-4), pp. 359-372.

[193] Spanos, J. T., Milman, M. H., and Mingori, D. L., 1992, "A New Algorithm for L2 Optimal Model Reduction", Automatica, **28**(5), pp. 897-909.

[194] Ferrante, A., Krajewski, W., Lepschy, A., and Viaro, U., 1999, "Convergent Algorithm for L2 Model Reduction", Automatica, **35**(1), pp. 75-79.

[195] Ungar, E. E., 1997, "Statistical Energy Analysis", Journal of Sound and Vibration, **31**(10), pp. 28-32.

[196] Rosenberg, R. C. and Zhou, T., 1988, "Power-Based Simplification of Dynamic System Models", *Proceedings of Advances in Design Automation*, American Soc of Mechanical Engineers (ASME), New York, NY, USA, **14**, pp. 487-492.

[197] Rosenberg, R. C. and Zhou, T., 1988, "Power-Based Model Insight", *Proceedings of Automated Modeling for Design, Nov 27-Dec 2 1988*, American Soc of Mechanical Engineers (ASME), New York, NY, USA, **8**, pp. 61-67.

[198] Louca, L. S. and Stein, J. L., 1998, "Physical Interpretation of Reduced Bond Graphs", *Proceedings of 2nd IMACS International Multiconference: Computational Engineering in Systems and Applications (CESA'98)*.

[199] Louca, L. S., Stein, J. L., and Hulbert, G. M., 1998, "A Physical-Based Model Reduction Metric with an Application to Vehicle Dynamics", *Proceedings of 4th IFAC Nonlinear Control Systems Design Symposium (NOLCOS 98)*, **3**, pp. 607-612.

[200] Louca, L. S. and Yildir, U. B., 2003, "Modeling and Reduction Techniques for Studies of Integrated Hybrid Vehicle Systems", *Proceedings of 4th International Symposium on Mathematical Modeling*.

[201] Rideout, D. G., Stein, J. L., and Louca, L. S., 2004, "Systematic Model Decoupling through Assessment of Power-Conserving Constraints - an Engine Dynamics Case Study", *Proceedings of 2004 ASME International Mechanical Engineering Congress and Exposition*, American Society of Mechanical Engineers, New York, NY, **2**.

[202] Fathy, H. K. and Stein, J. L., 2005, "Fundamental Concordances between Balanced Truncation and Activity-Based Model Reduction", *Proceedings of IMAACA '05, Bond Graph Methods for Dynamical Systems*, A. Bruzzone, G. Dauphin-Tanguy, C. Frydman, and S. Junco, eds., pp. 109-116.

[203] Ye, Y. and Youcef-Toumi, K., 1999, "Model Reduction in the Physical Domain", *Proceedings of American Control Conference*, Institute of Electrical and Electronics Engineers Inc., Piscataway, NJ, USA, **6**, pp. 4486-4490.

[204] Zames, G. and El-Sakkary, A., 1982, "Uncertainty in Unstable Systems: The Gap Metric", *Proceedings of Eighth Triennial World Congress of the International Federation of Automatic Control*, Pergamon, pp. 149-152.

[205] El-Sakkary, A. K., 1985, "The Gap Metric: Robustness of Stabilization of Feedback Systems", IEEE Transactions on Automatic Control, **AC-30**(3), pp. 240-247.

[206] Gough, N. E. and El-Sakkary, A. K., 1985, "Basic Procedure for Model Comparison Using the Gap Metric", International Journal of Systems Science, **16**(5), pp. 589-604.

[207] Georgiou, T. T., 1988, "On the Computation of the Gap Metric", Systems & Control Letters, **11**(4), pp. 253-257.

[208] Georgiou, T. T. and Smith, M. C., 1990, "Optimal Robustness in the Gap Metric", IEEE Transactions on Automatic Control, **35**(6), pp. 673-686.

[209] Chang, S. Y., Carlson, C. R., and Gerdes, J. C., 2001, "A Lyapunov Function Approach to Energy Based Model Reduction", *Proceedings of 2001 ASME International Mechanical Engineering Congress and Exposition*, American Society of Mechanical Engineers, **2**, pp. 1315-1322.

[210] Kübler, R. and Schiehlen, W., 2000, "Modular Simulation in Multibody System Dynamics", Multibody System Dynamics, **4**(2-3), pp. 107-127.

[211] Orlandea, N., Chace, M. A., and Calahan, D. A., 1977, "A Sparsity-Oriented Approach to the Dynamic Analysis and Design of Mechanical Systems - Part 1", **99 Ser B**(3), pp. 773-779.

[212] Assanis, D., Bryzik, W., Chalhoub, N., Filipi, Z., Henein, N., Jung, D., Liu, X., Louca, L., Moskwa, J., Munns, S., Overholt, J., Papalambros, P., Riley, S., Rubin, Z., Sendur, P., Stein, J., and Zhang, G., 1999, "Integration and Use of Diesel Engine, Driveline and Vehicle Dynamics Models for Heavy Duty Truck Simulation", 1999-01-0970, *Proceedings of 1999 SAE Congress*, SAE.

[213] Khazaka, R., Gunupudi, P. K., and Nakhla, M. S., 2000, "Efficient Sensitivity Analysis of Transmission-Line Networks Using Model-Reduction Techniques", IEEE Transactions on Microwave Theory and Techniques, **48**(12), pp. 2345-2351.

[214] Lin, R. M. and Lim, M. K., 1995, "Structural Sensitivity Analysis Via Reduced-Order Analytical Model", Computer Methods in Applied Mechanics and Engineering, **121**(1-4), pp. 345-359.

[215] Bladh, R., Castanier, M. P., and Pierre, C., 1999, "Reduced Order Modeling and Vibration Analysis of Mistuned Bladed Disk Assemblies with Shrouds", Journal of Engineering for Gas Turbines and Power, **121**(3), pp. 515-522.

[216] Sohns, B. J., Allison, J., Fathy, H. K., and Stein, J. L., 2006, "Accelerating Parameterizaton of Large-Scale Dynamic Models through the Use of Activity Analysis", *Proceedings of 2006 ASME International Mechanical Engineering Congress and Exposition*, ASME, New York, NY.

[217] Kozaki, T., Mori, H., Fathy, H. K., and Gopalswamy, S., 2004, "Balancing the Speed and Fidelity of Automotive Powertrain Models through Surrogation", *Proceedings of 2004 ASME International Mechanical Engineering Congress and Exposition*, ASME, New York, NY, **73**, pp. 249-258.

[218] Felgner, F., Merz, R., and Litz, L., 2006, "Modular Modelling of Thermal Building Behaviour Using Modelica", Mathematical and Computer Modelling of Dynamical Systems, **12**(1), pp. 35-49.

[219] Rinderle, J. R. and Subramaniam, B. L., 1991, "Automated Bond Graph Modeling and Simplification to Support Design", *Proceedings of Winter Annual Meeting of the American Society of Mechanical Engineers*, ASME, New York, NY, USA, **34**, pp. 45-68.

[220] 20-Sim, v3.6, 2005, Controllab Products B.V., Enschede, The Netherlands.

[221] Kane, T. R. and Wang, C. F., 1965, "On the Derivation of Equations of Motion", Journal of the Society for Industrial and Applied Mathematics, **13**(2), pp. 487-492.

[222] Kane, T. R. and Levinson, D. A., 1980, "Formulation of Equations of Motion for Complex Spacecraft", Journal of Guidance and Control, **3**(2), pp. 99-112.

[223] Wehage, R. A. and Haug, E. J., 1982, "Generalized Coordinate Partitioning for Dimension Reduction in Analysis of Constrained Dynamic Systems", Journal of Mechanical Design, **104**(1), pp. 247-255.

[224] Walton, W. C. and Steeves, E. C., 1969, "A New Matrix Theorem and Its Application for Establishing Independent Coordinates for Complex Dynamical Systems with Constraints", TR-326, NASA Technical Report.

[225] Kim, S. S. and Vanderploeg, M. J., 1986, "QR Decomposition for State Space Representation of Constrained Mechanical Dynamic Systems", Journal of Mechanisms, Transmissions, and Automation in Design, **108**(2), pp. 183-188.

[226] Singh, R. P. and Likins, P. W., 1985, "Singular Value Decomposition for Constrained Dynamical Systems", Journal of Applied Mechanics, **52**(4), pp. 943-948.

[227] Mani, N. K., Haug, E. J., and Atkinson, K. E., 1985, "Application of Singular Value Decomposition for Analysis of Mechanical System Dynamics", Journal of Mechanisms, Transmissions, and Automation in Design, **107**(1), pp. 82-87.

[228] Wehage, R. A. and Loh, W.-Y., 1993, "Application of Svd to Independent Variable Definition in Constrained Mechanical Systems", *Proceedings of ASME Winter Annual Meeting*, ASME, New York, NY, USA, **52**, pp. 71-79.

[229] Liang, C. G. and Lance, G. M., 1987, "A Differentiable Null Space Method for Constrained Dynamic Analysis", Journal of Mechanisms, Transmissions, and Automation in Design, **109**(3), pp. 405-411.

[230] Ider, S. K. and Amirouche, F. M. L., 1988, "Coordinate Reduction in the Dynamics of Constrained Multibody Systems-a New Approach", Journal of Applied Mechanics, **55**(4), pp. 899-904.

[231] Agrawal, O. P. and Saigal, S., 1989, "Dynamic Analysis of Multi-Body Systems Using Tangent Coordinates", Computers and Structures, **31**(3), pp. 349-355.

[232] Kurdila, A., Papastavridis, J. G., and Kamat, M. P., 1990, "Role of Maggi's Equations in Computational Methods for Constrained Multibody Systems", Journal of Guidance, Control, and Dynamics, **13**(1), pp. 113-120.

[233] Sleight, R. and Agrawal, S. K., 2004, "Dynamic Model of a Four-Wheel-Drive HMMWV", *Proceedings of ASME Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, ASME, **2 B**, pp. 1183-1191.

[234] Pacejka, H., B., 2002, *Tire and Vehicle Dynamics*, SAE, Warrendale, PA.

[235] Aardema, J., 1998, "Failure Analysis of the Lower Rear Ball Joint on the HMMWV", Technical Report, RD&E Center.

[236] DADS, v9.0, 1998, Computer Aided Design Software, Inc., Oakland, Iowa.

[237] Kittirungsi, B., Fathy, H. K., and Stein, J. L., 2006, "An Efficient Scaling Methodology for Dynamic Models Using Dimensional and Activity Analyses",

Proceedings of 2006 ASME International Mechanical Engineering Congress and Exposition (IMECE 2006), American Society of Mechanical Engineers, New York, NY 10016-5990, United States.