Frequency and time simulation of squeal instabilities. Application to the design of industrial automotive brakes.

Guillaume Vermot Des Roches

ECOLE CENTRALE DES ARTS ET MANUFACTURES

THÈSE

présentée par

Guillaume VERMOT DES ROCHES

pour l'obtention du

GRADE DE DOCTEUR

Spécialité : **Mécanique**

Frequency and time simulation of squeal instabilities

Application to the design of industrial automotive brakes

soutenue le 27/01/2011 devant un jury composé de

Président :	M. Daniel RIXEN	(TU Delft)
Rapporteurs :	M. Michel RAOUS	(LMA – Marseille)
	M. Jean-Jacques SINOU	(LTDS – Lyon)
Examinateur :	M. Remi LEMAIRE	(Robert Bosch France (SAS) – Drancy)
Co-Directeur de thèse:	M. Denis AUBRY	(MSS-MAT – Paris)
Co-Directeur de thèse:	M. Etienne BALMES	(SDTools – Paris)
Invité :	M. Xavier LORANG	(SNCF I&R – Paris)

Laboratoire de Mécanique des Sols, Structures et Matériaux – CNRS UMR 8579 – n° 2011 ECAP0009

Guillaume Vermot des Roches
Application Engineer
SDTools
151, Boulevard de l'hôpital, 75013, Paris
France

Phone: +33 1 44 24 63 50
e-mail: vermot@sdtools.com
www.sdtools.com

Abstract

Brake squeal is a common noise problem encountered in the automotive industry. Higher friction coefficients and weight reduction recently led to higher vibration levels in the audible frequency range. This quality issue becomes economic due to penalties imposed to the brake supplier although no robust design method exists. The industrial practice thus relies on costly prototyping and adjustment phases. The evolution of computational power allows computation of large mechanical assemblies, but non-linear time simulations generally remain out of reach. In this context, the thesis objective is to provide numerical tools for squeal resolution at early design stages.

Parameterized reduction methods are developed, using system real modes as Rayleigh-Ritz vectors, and allow very compact reduced models with exact real modes. The proposed Component Mode Tuning method uses the components free/free modes as explicit degrees of freedom. This allows very quick sensitivity computation and reanalyzes of an assembly as function of local component-wise parameters. Non-linear time simulations are made possible through two ingredients. A modified non-linear implicit Newmark scheme and a fixed Jacobian are adapted for contact vibrations. The brake is reduced keeping a superelement with exact real modes and a local non-linear finite element model in the vicinity of the pad/disc interaction.

A set of design tools is illustrated for a full industrial brake model. First, instant stability computations and complex mode trajectories are studied. Modal interactions and non-linear phenomena inside the limit cycles are thus well understood. Time/frequency correlations are performed using transient modal identification and space-time decomposition. A time domain modal damping model is also shown to be very useful. The modification of a critical component for squeal resolution is finally tested and validated.

Keywords: structural dynamics, self excited vibrations, brake squeal, contact-friction instability, time simulation, damping, component modes, parametric model reduction, substructuring, non-linear analysis, design.

Résumé

Le crissement de frein est une nuisance sonore classique dans l'automobile. L'augmentation des coefficients de friction et la réduction de la masse mènent aujourd'hui à de hauts niveaux vibratoires dans les fréquences auditives, et ces problèmes de qualité se traduisent par des pénalités économiques aux équipementiers, bien qu'il n'existe pas de méthode robuste de conception. La pratique industrielle repose donc sur de coûteuses phases de prototypage et d'ajustement. L'évolution de la puissance de calcul permet le calcul de grands assemblages mécaniques mais les études vibratoires non-linéaires restent généralement hors de portée. Dans ce contexte, l'objectif de la thèse est de fournir, dès les phases de conception, des outils de conception numérique d'aide à la résolution du crissement.

Une méthode de réduction paramétrée utilisant comme base de Rayleigh-Ritz les modes réels du système assemblé permet la génération de modèles réduits très compacts, avec modes réels exacts. La méthode proposée d'ajustement des modes de composants utilise les modes libres de composants comme degrés de liberté explicites. L'étude des sensibilités et la réanalyse d'un assemblage en fonction de modifications à l'échelle d'un composant deviennent possibles. Les études temporelles non-linéaires sont rendues possibles par deux développements. Un schéma de Newmark non-linéaire modifié et un Jacobien fixe adapté aux vibrations de contact sont introduits. Le frein est réduit en un superélément avec modes réels exacts et une zone non réduite au niveau du contact.

Un ensemble d'outils de conception est illustré sur un modèle industriel de frein. La stabilité instantanée et les trajectoires de modes complexes sont étudiées. Les interactions modales et les phénomènes non-linéaires au sein des cycles limites sont alors mieux compris. Des corrélations temps/fréquence sont obtenues par l'identification modale instantanée et une décomposition espace-temps. La grande utilité d'un modèle temporel d'amortissement modal est illustrée. Enfin, la modification d'un composant critique au crissement est testée et validée.

Mots-clés : dynamique des structures, vibrations auto-entretenues, crissement de frein, contact, instabilités de friction, simulation temporelle, amortissement, modes de composant, réduction de modèles paramétriques, sous-structuration, analyse non-linéaire, conception.

Contents

1

Introduction

1.1 Industrial stakes, the squeal problem

The ability to slow down or to stop is a necessary control feature of all vehicles. The concept of braking thus requires the implementation of an energy dissipation mechanism associated with forces resistive to the movement. Several braking means exist, using friction, aerodynamic resistance, power source constraining (engine brakes), *etc.* For car, train and plane applications, service braking is commonly performed by friction dissipation.

Friction is a classical source of noise. Adnan Akay [1] provides a comprehensive study of the *Acoustics of friction*, which can constitute a great introduction to any work treating this subject. Brake noise generates critical environmental disturbances. Akay notes for example that as early as in 1930, brake noise appeared in the top-10 noise problems surveys performed by New York City.

Brake noises belong to the class of friction induced vibration. They occur in many common systems for various configurations. The noise of a chalk on a blackboard, or the sound of a door opening are very classical examples.

Friction acoustics can however be pleasant to the human ear. The violin in particular is a canonical example, where stick-slip oscillations occur when the bow enters in contact with the string, responsible for the sound. The analysis will then not focus on eliminating the noise, but on controlling the note produced. A compilation of self excited oscillation examples with friction has been performed for example by Popp *et al.* [2].

For Bosch[1], the industrial partner of this project, recent technological developments are creating more and more noisy systems. Damping levels have been reduced due to the ban on asbestos for sanitary reasons. This material was used in the friction linings and had a viscoelastic effect, which is not reproduced with the new material formulations. Systems are more unstable due to the increase of the friction coefficient. More performance from the brake systems is demanded in conjunction with the creation of heavier cars, and the increase of safety requirements. More system modes occur nowadays in the audible frequency range (under 16 kHz) due to the weight reduction performed. Brakes are indeed non-suspended masses whose effect is critical on chassis balance.

The disturbance first has an environmental aspect, since it generates noise pollution. Noise occurs mostly at low speed and low braking pressure levels. It mainly corresponds to urban uses, where interactions with habitations areas and pedestrians are the highest. At low speed other noises are also lower, so that brake noises are then very clear to the driver. Thus, although brake noise generation is usually not a sign of performance degradation, it is typically perceived as a defect that must be corrected and generates car returns.

Warranty claims and loss of perceived quality are clearly negative for the brand image of car manufacturers. Economic penalties are thus imposed to brake suppliers in cases of noisy brakes. But although control of noise occurrences is a major industrial problem, no robust design tool exists at the moment. Brake noise is still handled using prototyping/validation phases, which are costly and require a lot of time.

The noise generated by a brake system can range from 10Hz to 16kHz. Every car maker has a nomenclature of its own to classify noises. Akay [1] gives an example of classification, reproduced in figure 1.1.

A lot of noise classes have in common their generation mechanism; their expression can differ in terms of modulation, evolution patterns, or in noise durability.

The lower frequency noises can be of high amplitude and be transmitted to other car components. In particular Judder, which is mainly due to disc thickness variations, can disturb the steering wheel and alter the braking performance. Squeal noise is higher pitched and can attain 120 dB in the system vicinity.

[1]Robert Bosch France (SAS) Chassis Systems Brakes, Drancy

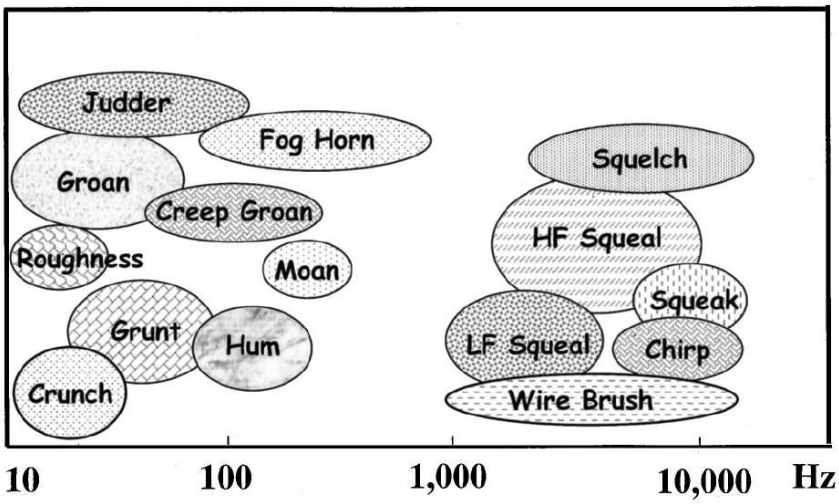


Figure 1.1: Brake noise classification by spectral contents, as proposed by Akay [1]

Brake squeal has therefore long been a very active field of research, from the work of North *et al.* in the 70's [3] to the most recent. The work of Vola [4] under the direction of Raous, Moirot [5] to Lorang [6] under the direction of Nguyen, Massi [7], Linck [8] and Meziane [9] under the direction of Baillet, Fritz [10], and Coudeyras [11] under the direction of Sinou are examples of recent (1998) to very recent (2009) PhD theses devoted to the subject.

The works by Akay [1], Ouyang *et al.* [12, 13], Kinkaid *et al.* [14], Fieldhouse [15], and Hoffman [16] are also representative of advances in the subject.

Manufacturing companies are eventually very eager to publish on their advances, in the SAE conferences for example [17]. It can be noted that Bosch is a very active player and launched numerous studies based on a simplified beam-on-disc system. Its objective is of course design oriented. One would retain [18] by Tan *et al.* as an ambitious paper treating of design processes against brake squeal. The design concepts that will be developed in the present thesis to greater extents can find its root in [18], regarding the approach Bosch employs to treat its noise issues.

The thesis has been performed at SDTools, in partnership with Bosch. The stakes for SDTools in this context are then different. The proposed application is new, as the time simulation of large industrial models will be performed. The implementation of such procedure and an increase in knowledge of time integration and non-linear methods is thus needed.

The design application is also novel, proposing to keep component modes as explicit degrees of freedom at the assembly level. The proof of concept and finding exploitable applications is then the priority.

1.2 An industrial brake model

The present work will focus on a case of study provided by Bosch. The full brake assembly is shown in figure 1.2. It illustrates a disc brake with floating piston, representative of the brakes found in classical berlines produced by Bosch customers such as PSA, Fiat, VAG, *etc.*

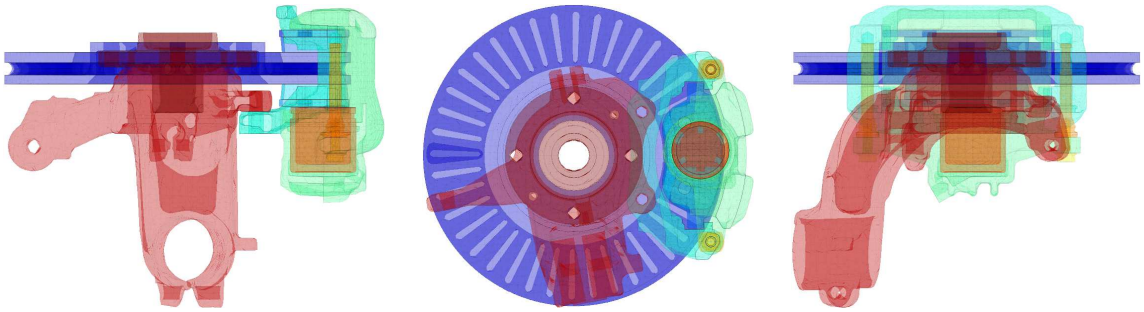


Figure 1.2: Global view of the industrial brake model provided by Bosch

The original brake design is traditionally studied for Noise Vibration and Harshness (NVH) applications by the simulation team of Bosch Chassis Brake Systems. The brake targeted is free meshed using quadratic 3D 10 nodes tetrahedrons, thus yielding a model of approximately 600,000 Degrees Of Freedom (DOF).

The model is an assembly of eight components comprising a rotor assembly and a stator assembly. The rotor part presented in figure 1.3 is made of a *disc*, and a *hub*. In practice the disc is fixed on the wheel and constitutes the part on which the braking pressure is effective.

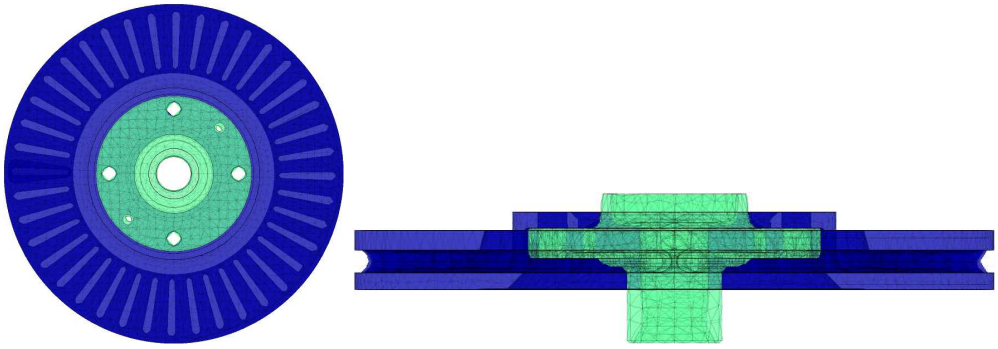
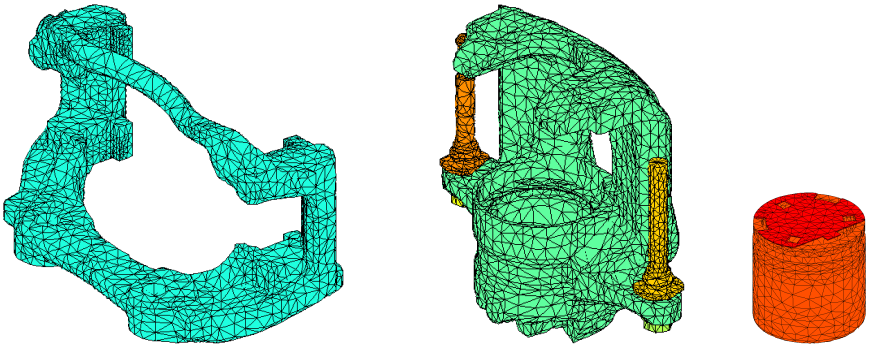
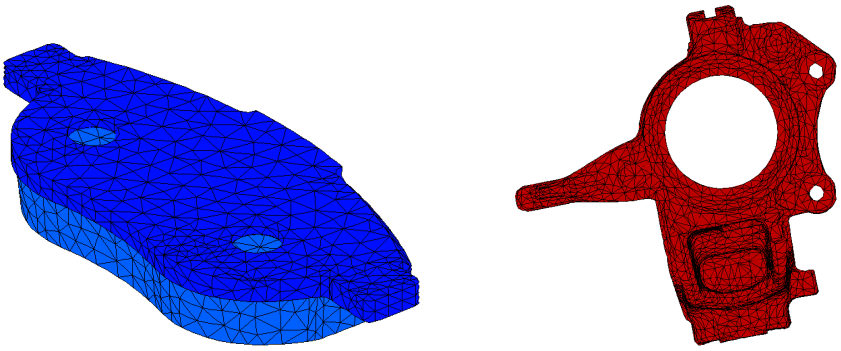


Figure 1.3: Rotor part of the brake system, disc (blue) and hub (green)

The stator realizes the link to the chassis, via a *knuckle*, figure 1.4e, which also supports the brake mechanism, constituted of an anchor, figure 1.4a, caliper, figure 1.4b, piston, figure 1.4c and 2 pads, figure 1.4d.



(a) Anchor (b) Caliper (c) Piston



(d) Pad (e) Knuckle

Figure 1.4: Stator components

The brake mechanism, shown in figure 1.5, applies the braking pressure to the disc, linked to the knuckle through the hub. In practice, the rotor is mounted to a bearing, modeled by equivalent stiffness coupling.

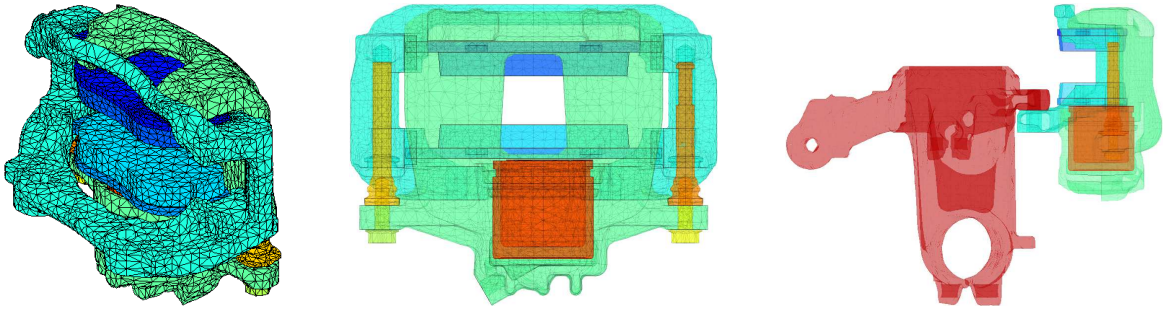
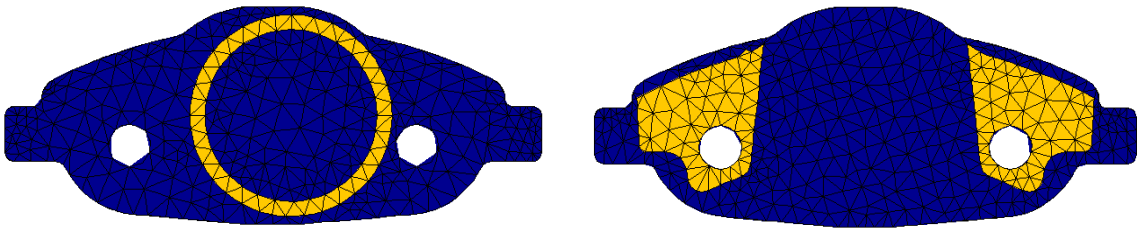


Figure 1.5: Stator part of the brake system

A single piston is used for the floating type of mechanism. The braking pressure from the pedal is transmitted to the piston through a hydraulic circuit. The piston, translating in the caliper, pushes the inner pad against the disc. Once the inner pad is in contact with the disc the piston reaction pushes the caliper away, inducing its translation in the anchor, and pressure on the outer pad, until equilibrium is found.

The floating piston mechanism is widespread as it is cheaper to produce than multi-piston systems, and provides sufficient performance for most cars in the market. The pad loading resulting from this technology is unsymmetric. The inner pad (piston side) is loaded at its center, while the outer pad (caliper side) is loaded by the caliper fingers at both ends. This pattern is illustrated in figure 1.6, and will be the key to understanding some of the phenomena analyzed in the application.



(a) Inner pad (piston side) (b) Outer pad (caliper side)

Figure 1.6: Braking pressure transmission areas on the pads

The mechanism is rather complex. Interactions between components showing non negligible displacement are modeled using contact-friction laws. This concerns the relative motion of the pads in the anchor, translation of the caliper in the anchor, movement of the piston in the caliper, motion of the pads in the caliper and motion of the pad against the disc. All these contact interactions are illustrated in figure 1.7, which represent the potential contact areas between all components in the assembly. The contact models used are tabulated pressure-gap relationships, defining penalized contact formulations. The contact laws are identified experimentally and take into account non modeled parts such as bushes, pad springs, *etc.*

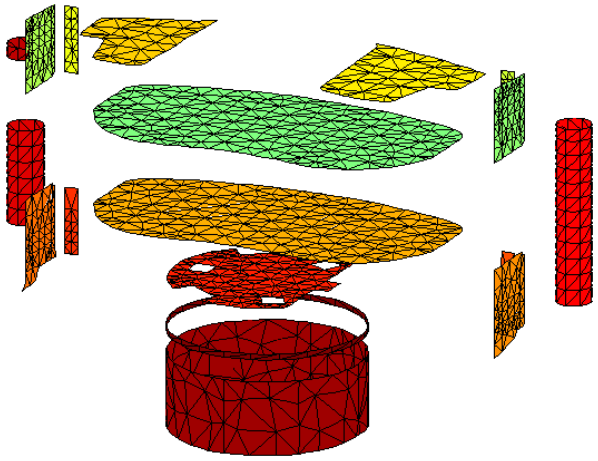


Figure 1.7: Contact elements of the brake system

Fixed boundary conditions are applied to the knuckle fixations on the Chassis. The simulations suppose a steady state rotation of the disc (imposed by the wheel) and thus require braking torque compensation. Bolts, which fix the wheel are thus replaced by grounded elastic coupling stiffness coupled to the holes as presented in figure 1.8.

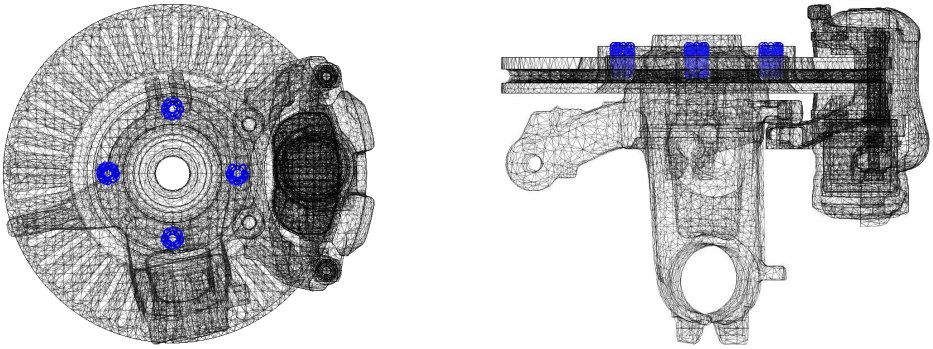


Figure 1.8: Bolt fixation DOF of the rotor with the wheel (o)

The working pressure point of most interest for Bosch on this specific design was 12 Bar. The static states, presented in figure 1.9, are provided by Bosch, and computed with ABAQUS using large displacement non linear static analyses. Due to the friction forces, and the disc rotation, the outer pad bends at the disc entry and generates a transverse displacement of the caliper and anchor towards the area where the disc enters under the pad. This generates a well known arching positioning of the pads.

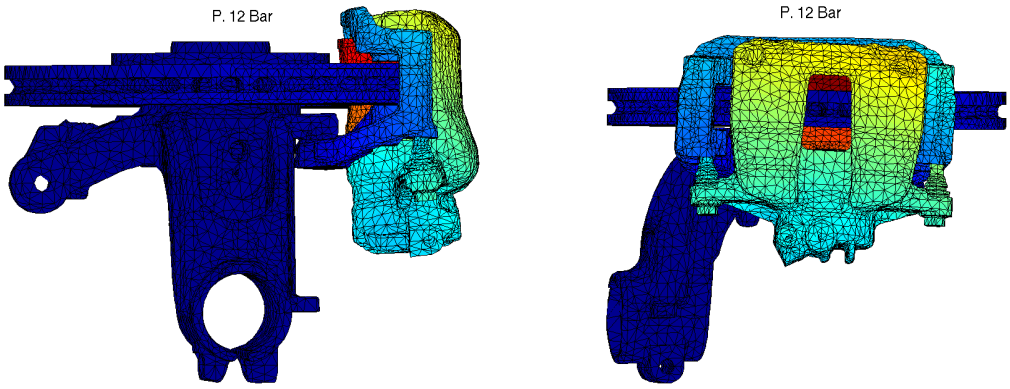


Figure 1.9: Static deformation of the industrial brake model for a braking pressure of 12 Bar

For visualization purposes an observation mesh, shown in figure 1.10, is placed on the disc. The objective here is to produce results that can be compared to experimental data in the form of Operational Deflection Shapes (ODS), that are regularly obtained by Bosch. A validation of the results presented can then be made by comparing test and simulation vibration levels on the ODS mesh.

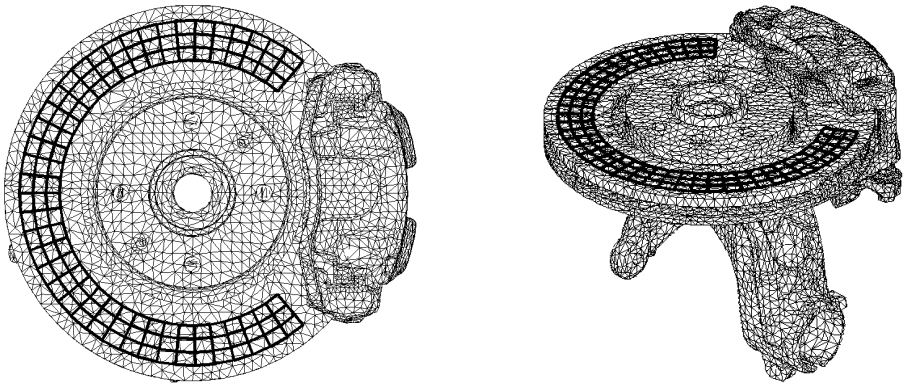


Figure 1.10: Observation mesh placed on the disc for Operational Deflection Shapes (ODS)

A lot of experimental characterizations are performed by Bosch, from component updating to squeal reproduction. The installations allow controlling several parameters, such as ambient temperature and hygrometry. Tests are realized for systems alone and for systems mounted in car parts or even full cars. A sample squeal test setup is presented in figure 1.11, where the disc model presented above is mounted.

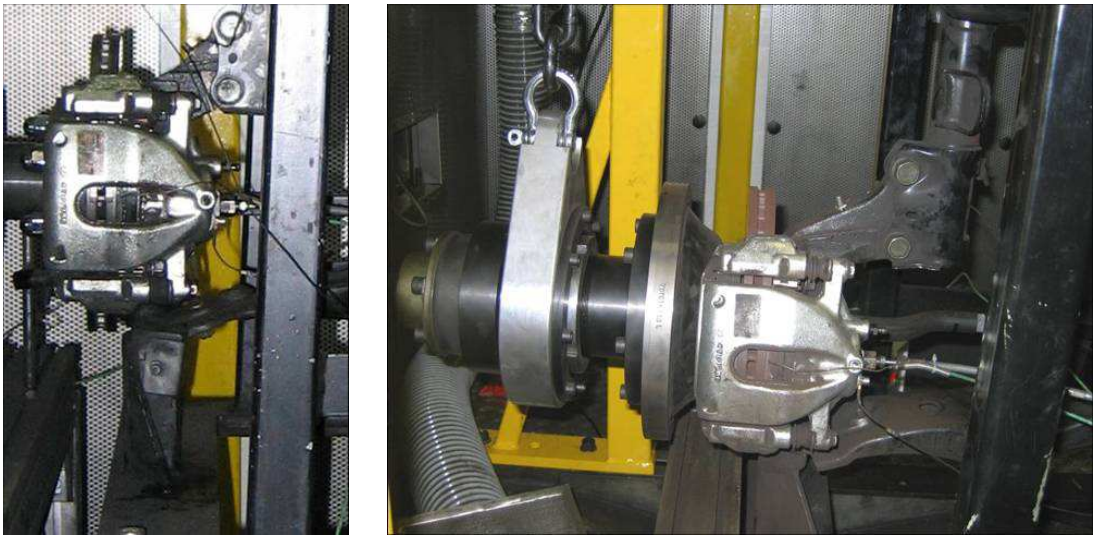


Figure 1.11: A squeal test bench at Bosch facilities

Squeal detection is performed surveying noise levels in the test cabin. If a specific noise level is passed, a squeal occurrence is registered, and the noise signal is recorded. FFT are performed for each occurrence, producing global performance results for the brake.

Squeal occurrences are represented by their noise level against frequency, producing graph reports like the one presented in figure 1.12. The brake of interest is then noisy, showing two clear noise frequencies at 4kHz and 10 kHz. A refined observation concluded in particular that the 4kHz noise occurs mainly at 12 Bar. The reproduction of such pattern is thus a target of the thesis.

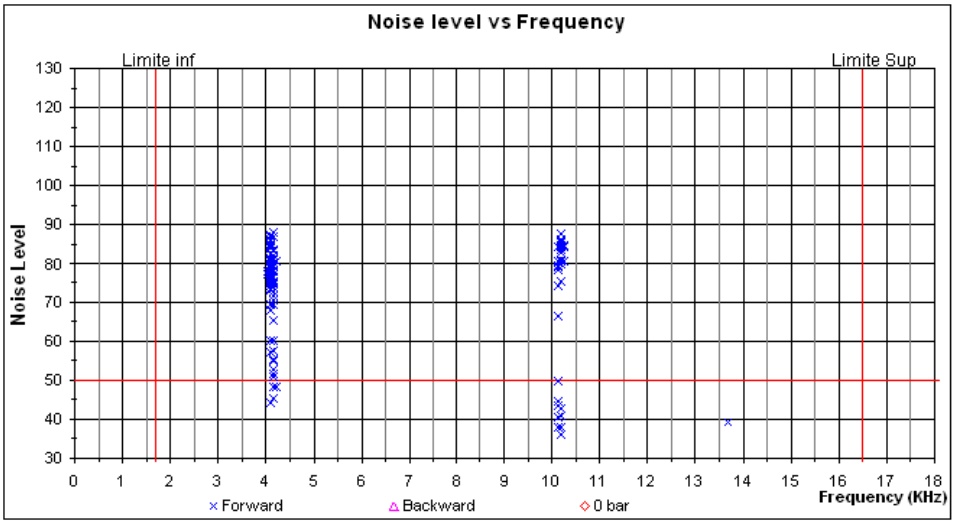


Figure 1.12: Experimental squeal results for the modeled brake

In the course of the thesis, a simplified brake model, presented in figure 1.13, was used for feasibility studies and proof of concept. It will be used for illustrations. Its material properties are presented in table 1.1 and provide representative values of the true application whose characteristics are confidential.

No in depth analysis of this simple brake will be presented in the following. to illustrate the relative stiffness differences happening in such brake assembly.

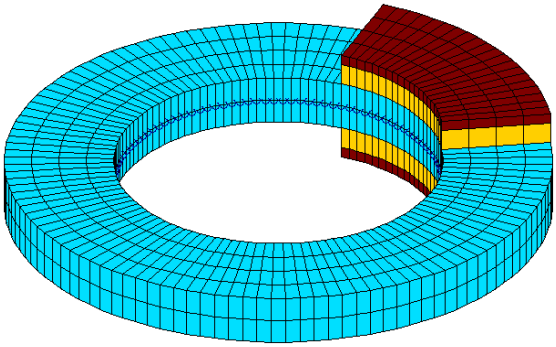


Figure 1.13: Simplified brake model used for code validations

Component	Modulus	Poisson coefficient	Density
disc	$119GPa$	0.25	$7100kg/m^3$
backplate	$210GPa$	0.3	$7800kg/m^3$
lining	$E_{xx} = E_{yy} = 7.6GPa$ $E_{zz} = 2.8GPa$ $G_{xz} = G_{yz} = 1GPa$ $G_{xy} = 3.8GPa$	$\nu_{xz} = \nu_{yz} = 0.25$ $\nu_{xy} = 0.125$	$2620kg/m^3$

Table 1.1: Material values used for the simplified brake model

The pads are constituted of two materials, a metallic *backplate* and a *lining* or *friction material* coming into contact with the disc. The lining material must be much softer than the disc and feature great braking capacity while enduring difficult mechanical, thermal, chemical constraints. In practice the lining is a composite fretted material. It has an orthotropic elastic behavior for which the transverse Young modulus (E_{zz}) will be involved for pad compression characterizations. A large Young modulus ratio can then be observed at the pad/disc interface and non-quasistatic wave propagation within the lining can occur in the frequency band of interest.

1.3 Objectives, contributions and thesis outline

Integrated design requires an extensive use of simulation to cut down on prototyping and testing time, and help focus on more promising designs. Although testing is expected to remain, simulation also provides a much deeper insight into the systems. In the case of braking, access to detailed information about contact pressures is of particular interest. The ability to run many design variants and simulation *replays* is another advantage.

To become useful, simulations must however be validated. Validation is the process in which one establishes the ability of simulations to reproduce all important behavior seen in test. This ability is determined with tolerances and an optimized compromise between the richness of the models simulated – their level of complexity – and computation times has to be found. Depending on the computational power and model size, some accuracy often has to be sacrificed to obtain a result in a relevant time. Such trade-offs are at the heart of the present research. Issues that need to be addressed are the ability to simulate full brake assemblies over long periods of time.

Transient simulation of the considered brake model requires a time step under $10^{-6}s$. As 100ms simulations require 100,000 time steps, for a 600,000 DOF model, the time simulation result would lead to a 500 GB output. The CPU time to compute residue and linear resolutions would also range in months. A reduction technique is therefore required to limit the model size. Novel Component Mode Synthesis techniques developed by Balmes [19, 20] were used to generate very compact systems with exact reduced real modes, here yielding reduced models with 30,000 DOF for transient analysis. The application of these methods to squeal problems is detailed in section 2.2 and publication [21] and is a **first contribution** of this work.

The development of a product (here a mechanical assembly) is conceptualized by a V-cycle, as the one shown in figure 1.14. From the concept to the operational model, a design phase first creates local components with proper characteristics.

A second phase deals with the integration of the designed components altogether and evaluate the assembly behavior, for which specifications are usually expressed. When specifications are not matched a loop is performed in the cycle where, from global observations at the assembly level, local design modifications are made and validations performed again.

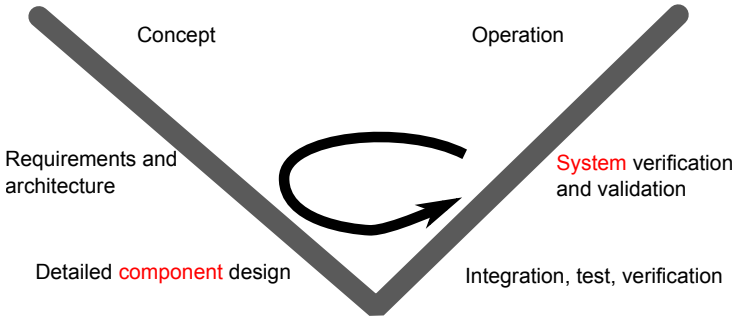


Figure 1.14: V-model for product development

In vibration applications, the prediction of operation is often only relevant when using fully assembled system models which correspond to the end of the V cycle. Understanding how to modify components to achieve the desired changes on the system behavior is thus critical.

The classical approach is to propose physical changes, perform sensitivity computations and full solutions. An industrial example is for example shown in [22]. Such computations are costly and parametric reduced models [23] typically loose information about components. Reconnecting, with an acceptable numerical and data input cost, the expression of changes at the assembly level to the design parameters at the component level is thus a major challenge.

In experimental modal analysis, *Structural Dynamics Modification* methods, see [24] for

example, consider a component known through testing and modifications that are modeled. Recent developments by Corus [25] and Groult [26] under the direction of Balmes have led to the more explicit use of the main component modes as degree of freedom for the system model. Combining this idea with parametric reduction methods is the base of the **second contribution** of this work, the **Component Mode Tuning** method, see section 2.3 and publication [27].

In the proposed process, exact system modes are needed and efficient method for their computation is a key requirement. Automated Multi Level Solvers (AMLS), developed in particular by Kaplan [28] and Bennighof [29] for FRF computation and eigenvalue resolution, are the fastest solvers currently available. These methods are based on domain decomposition and parallel resolution and allow computation of systems of over a few million DOF on workstations in a reasonable time. An implementation of such algorithms is discussed in section 2.4, with the purpose of highlighting important trade-offs for future improvements in the implementation.

For realistic applications in dynamics, the response of complex mechanical systems is usually needed. Such systems are *assemblies* of separate *components* in *interaction* with each other. Vibration studies at the scale of a component, assuming elastic materials in small deformations, are nowadays very common, well mastered and computationally accessible. The problem has thus naturally shifted towards modeling of component interfaces, with sliding and sticking contact behavior being critical in brake assemblies.

Industrial meshing constraints often lead to incompatible and possibly coarse discretization of surfaces in contact. Subspace methods to avoid classical irregularities found in such setting are thus explored section 2.5 and publication [30, 31]. This **third contribution** was to provide a new measure of compatibility between two finite element meshed interfaces and how this can be used to limit locking issues found with the classical over-integration approach.

Since the focus is placed on simulating large industrial assemblies, contact-friction modeling is treated rather conventionally in chapter 3. Section 3.2 thus starts with an overview of the common contact-friction models found in the literature. The traditional penalization and Lagrange differences in strategy are discussed.

Section section 3.3 then addresses classical non-linear instability analyses. The simulation of non-linearities is complex and should be performed in the time domain. Their effect can then be assessed precisely. Due to the difficulty of simulating large models in the time domain, so-called non-linear methods have been developed to compute approached responses from an unstable system. A brief bibliography of such methods is presented, followed by a summary of choices made within the thesis.

Methods classically applied in industry are stability computations based on the tangent state of a system to a design working point. Although this does not provide information about the non-linear evolution of the system, it gives initial clues on the propensity of a system to be unstable. Such computations were therefore performed on the brake model and are presented in section 3.4. The **contribution** of the chapter (see publications [32, 33]), is thus to provide extensions of the stability method by using the CMT, to propose the use of complex modes trajectories as pseudo-cycles to enhance understanding, and to discuss the fact that the use of a static state for stability analysis is not the most relevant.

A major objective is to provide squeal cycles of real brake models, which cannot be obtained by stability analyses. *Brute force* was therefore employed through the implementation of a time simulation method adapted to reduced industrial brake models with local non-linearities. This is the objective of chapter 4. First an overview of the classical time integration schemes is presented in section 4.1. The implicit Newmark scheme was chosen, with adaptations presented in section 4.2. The **contribution** is associated with the use of a fixed Jacobian based on bilateral contact to perform the time integration. This feature solves classical convergence difficulties and allows reducing time computations for 30,000 DOF models from several thousands of hours to

around 12 hours. Time simulation results of the industrial brake are provided in section 4.3.

Dynamic modeling of contact friction poses time and space scales problems. Coarse meshes are typically unacceptable for spatial description of contact and friction. To resolve such scales issues in simulation, multi-scales approaches have been developed. Ben Dhia and Rateau [34] develop the so-called *Arlequin* method. This method is a multi-model method allowing coupling of multiple discretizations of specific parts of a component. The coupling strategy is then based on energy distribution schemes between the models. Combescure and Gravouil [35] developed space-time multi-scales approaches taking into account non-linearities. These methods are however limited in their ability to handle long time simulations of large models.

A classical approach is to observe contact and friction at a global scale where the locally simple contact/friction phenomena lead to very complex behavior with discontinuities, hysteresis, *etc.* For example, looking at the atomic scale, an impact between two bodies can be observed as a continuous interpenetration of atomic layers; while globally and coarsely, the impact can be seen as the displacement discontinuity of a solid. The work of Ouyang *et al.* [13] illustrates for example the difficulty to obtain realistic contact pressure distributions on simple brakes modeled in finite elements.

It is useful to note that this scale problem is critical experimentally. Indeed, the measurements become very difficult or even impossible when trying to identify the properties of two interacting interfaces. Since measurements cannot be intrusive – otherwise the interaction is altered – direct experiments cannot be performed. Besides, high-frequency excitation is difficult or impossible. Brake squeal goes up to 16kHz, which is far beyond most experimental means for large systems. Measurements by Massi [7] or Bureau *et al.* [36, 37] for contact-friction systems represent recent applications in this area.

In the proposed simulations, the time and space scales associated with non-linear contact and friction between the pad and the disc are very small and quite different from the relatively large scales associated with the full brake assembly, which only needs to be modeled up to 16 kHz. The reduction method, proposed in section 2.2.5, allows small time steps to achieve good representation of local transients and locally unreduced meshes for a good accuracy on wave propagating in the soft lining. These ingredients are quite critical for the confidence placed in the results.

With the proposed time simulation capability, a first question addressed is the correlation between frequency domain stability analyzes and time simulations. For this purpose, one introduces and uses in section 4.4 a *modal damping model* for time domain simulations, and the dual *modal sensors*. This **contribution** is also the object of publication [38].

Section 4.5 concludes the chapter by analyzing so-called limit cycles. These cycles provide indications of physical amplitudes but are typically not related to design tools. A space-time decomposition is thus proposed and used for analysis.

Chapter 5 eventually applies the concepts presented in the previous chapters to perform design analysis of industrial brake squeal. First, the mechanisms behind the squeal cycles are analyzed through modal damping and transient stability characterization. The transient component coupling evolution in a squeal cycle is then highlighted in section 5.1. From this information, the design study is focused at the component level in section 5.2. Relevant component design information are extracted in the frequency domain and validated in the time domain which suggests an improved design process for squeal applications.

2

Reduction methods for assemblies

Contents

2.1 Introduction

For most vibration applications, the system can be characterized by its normal modes. In the present case despite a working frequency band of several kHz, only a few hundred deformation shapes are necessary. Large models, typically with millions of DOF, are thus used to properly account for geometric detail necessary to model industrial products, while few modes account for dynamic detail. The difficulty addressed in this work is that the reduction must be adapted to account for spatial and time detail associated with non-linearities and allow design studies.

Rayleigh-Ritz reduction is the basis of the approach. One introduces a reduction basis $[T_R]_{N\times n}$, so that finite element DOF $\{q\}_{N\times 1}$ are written as the combination $\{q_R\}_{n\times 1}$ of the basis n Ritz vectors

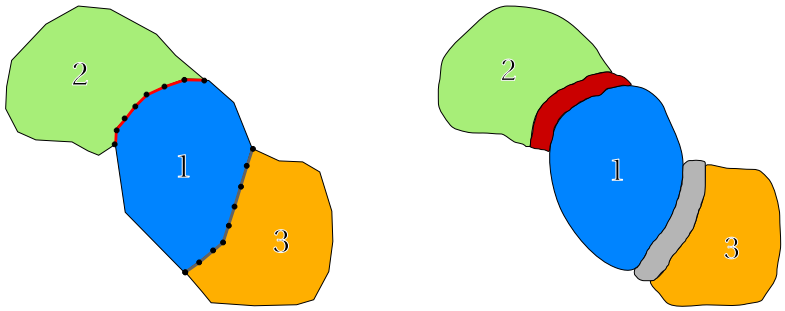
$$\{q\}_{N\times 1} = [T_R]_{N\times n}\{q_R\}_{n\times 1} \tag{2.1}$$

The reduced model DOF are not local finite element DOF expressing the displacement of a particular mesh point, but the contribution of a global shape in the response, hence the classical denomination of generalized DOF. The choice of the reduction basis must optimize the compromise between final size and precision. This is the topic of Component Mode Synthesis (CMS) or substructuring [39, 40] which has been studied since the 70's.

CMS deals with the coupling of components that interact with each other. Figure 2.1 illustrates the two main strategies, displacement or energy, used to account for the component coupling.

Displacement approaches, illustrated in figure 2.1a, assume the continuity of displacement and zero work of interface forces. This is commonly labeled as a *primal* method and is the classical framework of CMS. Section 2.2.1 discusses this approach and its main limitation associated with the treatment of large interfaces.

Energy methods, illustrated in figure 2.1b, uses elements to compute the energy associated with relative displacement of disjoint components. Although in general physical elements are preferable, penalty based formulations of the contact energy are also possible. In the present application, coupling is associated with displacement based formulations of contact pressure discussed in section 3.2.2. It should be noted that although this method uses loads to couple components, it does not correspond to *dual* methods discussed in [40].



(a) Displacement coupling (b) Energy coupling

Figure 2.1: Interface representation, coupling strategies

E. Balmes *et al.* published many papers and directed several theses on the optimization of reduction bases [41, 25, 26, 20, 42, 43, 44, 45, 23, 46, 47], for the Component Mode Synthesis method, expansion method and their extensions. It constitutes the theoretical basis of the methods implemented in SDT, and is part of the advanced design methods provided by SDTools.

From this work, three ideas were reused here. First, reduction of a component on the basis of exact system modes, previously used in [20] in particular, provides a mechanism to generate extremely compact models that have exact modes. This approach, discussed in section 2.2.2 to

2.2.5, allows accurate representation of system dynamics while retaining the ability to describe non-linear effects that are very local in space and time.

Second, brake components are usually analyzed and tested in their free/free configuration so that understanding the relation between the properties of these component modes and the system behavior is critical. The idea of using arbitrary component modes as DOF, present in [25, 26], is thus reused here.

The last idea is associated to the ability to use reduced models for parametric analysis [41, 20, 42, 23, 46]. Combined with the use of component modes as degrees of freedom, this leads to the Component Mode Tuning (CMT) method discussed in 2.3. This approach allows very precise reanalyzes, that is low cost predictions of a non nominal model using a reduced model generated in the nominal configuration. The use of component modes gives direct access to component information and easier iterations in the V-cycle presented in the introduction 1.3. Reanalysis of parametric models gives access to very fast approximations. A sample plate and the industrial brake are used for illustrations in section 2.3.

Since the proposed reduced model uses nominal system modes for their generation, high performance solvers for these modes are of interest. Extensions of CMS have been recently proposed by Bennighof and Kaplan [28, 29] with their Automated Multi Level Solvers (AMLS). The principles of such approaches, including decomposition and solution phases, are detailed in section 2.4 to optimize eigenvalue computations, the objective is a long term improvement of the eigenvalue solvers of SDT.

Finally in practical contact applications, models will often not be conforming. That is to say that interface DOF will not be coincident. The matching thus becomes non-trivial and associated issues are presented in section 2.5. In particular, the transmission of loads can lead to irregularities which must be handled, and a general framework extending the concept of $(1-\epsilon)$ compatibility, developed by Ben Dhia and Balmes [48], is presented.

2.2 Component Mode Synthesis concepts

2.2.1 Coupling by displacement continuity - Craig Bampton

The Craig-Bampton method [39] is the most widespread reduction method in industry. Its key ingredients are the explicit use of interface DOF and a reduction of components with no *a priori* global knowledge. The section details the underlying theory and illustrates the typical limitations associated with the size of interface matrices and the difficulty to improve accuracy beyond the nominal accuracy.

Each structure, namely 1 and 2 can be considered as having two separate type of DOF, *interface* DOF, denominated by the subscript i, and *complementary* DOF, denominated by the subscript c and representing the structure DOF outside the interface. As stated in the introduction 2.1, a substructure is here taken as a component. The displacement vector of component 1, q_1, can be described considering splitting DOF in $\{q_{i1}\}_{N_{1i}\times 1}$ and $\{q_{c1}\}_{N_{1c}\times 1}$. Structure 2 can be handled in the same way splitting q_2 into $\{q_{i2}\}_{N_{2i}\times 1}$ and $\{q_{c2}\}_{N_{2c}\times 1}$. The dynamic stiffness of both independent structures Z_i is written

$$\begin{cases} [Z_1] = \begin{bmatrix} Z_{1cc} & Z_{1ci} \\ Z_{1ic} & Z_{1ii} \end{bmatrix} \\ \\ [Z_2] = \begin{bmatrix} Z_{2cc} & Z_{2ci} \\ Z_{2ic} & Z_{2ii} \end{bmatrix} \end{cases} \tag{2.2}$$

The assembly itself is described by a displacement vector q comprising both component displacement vectors,

$$\{q\} = \left\{ \begin{array}{c} q_{c1} \\ q_{i1} \\ q_{i2} \\ q_{c2} \end{array} \right\}_{(N_{1c}+N_{1i}+N_{2i}+N_{2c})\times 1} \tag{2.3}$$

Component coupling is here expressed as a displacement continuity condition at the interface, which can be written

$$\{q_{i1}\}_{N_{1i}\times 1} = \{q_{i2}\}_{N_{2i}\times 1} = \{q_i\}_{N_i\times 1} \tag{2.4}$$

The coupling condition of equation (2.4) generates information redundancy in the displacement vector given in equation (2.3), which can be simplified, by introducing a single interface displacement vector q_i, defined in equation (2.4),

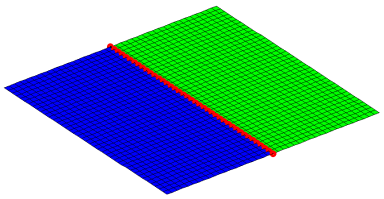
$$\{q\} = \left\{ \begin{array}{c} q_{c1} \\ q_{i} \\ q_{c2} \end{array} \right\}_{(N_{1c}+N_{i}+N_{2c})\times 1} \tag{2.5}$$

Using the simplified assembly displacement vector defined in equation (2.5), the coupling of components #1 and #2, whose dynamic stiffnesses are given in equation (2.2) can be written

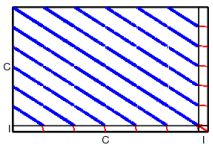
$$[Z]_{(N_{1c}+N_i+N_{2c})\times(N_{1c}+N_i+N_{2c})} = [Z_1] + [Z_2] = \left[\begin{array}{ccc} Z_{1cc} & Z_{1ci} & 0 \\ Z_{1ic} & Z_{1ii}+Z_{2ii} & Z_{2ic} \\ 0 & Z_{2ci} & Z_{2cc} \end{array} \right] \tag{2.6}$$

where all blocks are sparse matrices.

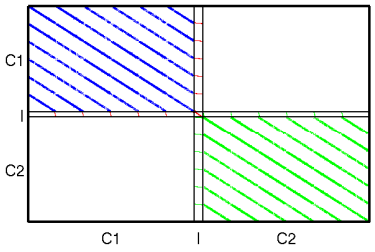
As an illustration, a sample steel plate, presented in figure 2.2a is chosen. The CMS is applied to the plate cut in two, so that both parts can be considered independently and the matrix topologies of equation (2.7) are exhibited in figure 2.2b. The assembled model (2.6) is shown in figure 2.2c.



(a) Plate topology



(b) Structure #1 topology



(c) Assembled stiffness matrix topology, NZ=307,216

Figure 2.2: CMS applied to a square plate cut in two. (I) for interface, (Ci) for complementary DOF of structures #1 and #2

If no external force is applied to the complementary DOF of both components, the response of the complementary DOF directly depends on the interface movement, so that it is possible to write a condensation basis for the stiffness matrix of structures 1 and 2

$$\left\{ \begin{array}{l} \left\{ \begin{array}{c} q_{c1} \\ q_i \end{array} \right\}_{(N_{1c}+N_i)\times 1} = \left[\begin{array}{c} -Z_{1cc}^{-1} Z_{1ci} \\ \mathcal{I} \end{array} \right] \{q_i\}_{N_i\times 1} \\ \\ \left\{ \begin{array}{c} q_{c2} \\ q_i \end{array} \right\}_{(N_i+N_{1c})\times 1} = \left[\begin{array}{c} -Z_{2cc}^{-1} Z_{2ci} \\ \mathcal{I} \end{array} \right] \{q_i\}_{N_i\times 1} \end{array} \right. \tag{2.7}$$

Under such assumption, the assembled dynamic stiffness matrix can be condensed on its interface DOF, writing

$$\left[Z_{ii_{N_i\times N_i}}\right] = \left[\hat{Z}_{1ii}\right] + \left[\hat{Z}_{2ii}\right] \tag{2.8}$$

where the hat matrices are the projection of the dynamic matrices on the condensation bases from equation (2.7),

$$\begin{cases} \left[\hat{Z}_{1ii}\right]_{N_i\times N_i} = Z_{1ii} - Z_{1ic}Z_{1cc}^{-1}Z_{1ci} \\ \left[\hat{Z}_{2ii}\right]_{N_i\times N_i} = Z_{2ii} - Z_{2ic}Z_{2cc}^{-1}Z_{2ci} \end{cases} \tag{2.9}$$

This formulation has limitations in terms of sparsity, as the condensed matrix becomes full. Above a few thousand interface DOF, the computational resources necessary to handle a full matrix is underperforming in comparison to a much larger sparse matrix.

A second limitation comes from the dynamic character of the condensation. The exact condensation is only available in the frequency domain and, for FRF computations, the computational cost for thousands of frequency points is excessive. In practice, the condensation is thus performed at zero frequency, $Z(\omega = 0) = K$. The shapes from the Guyan condensation, coupling the complementary DOF to the interface represent the response of the component to solicitations from each interface DOF. They are commonly named *constraint modes*.

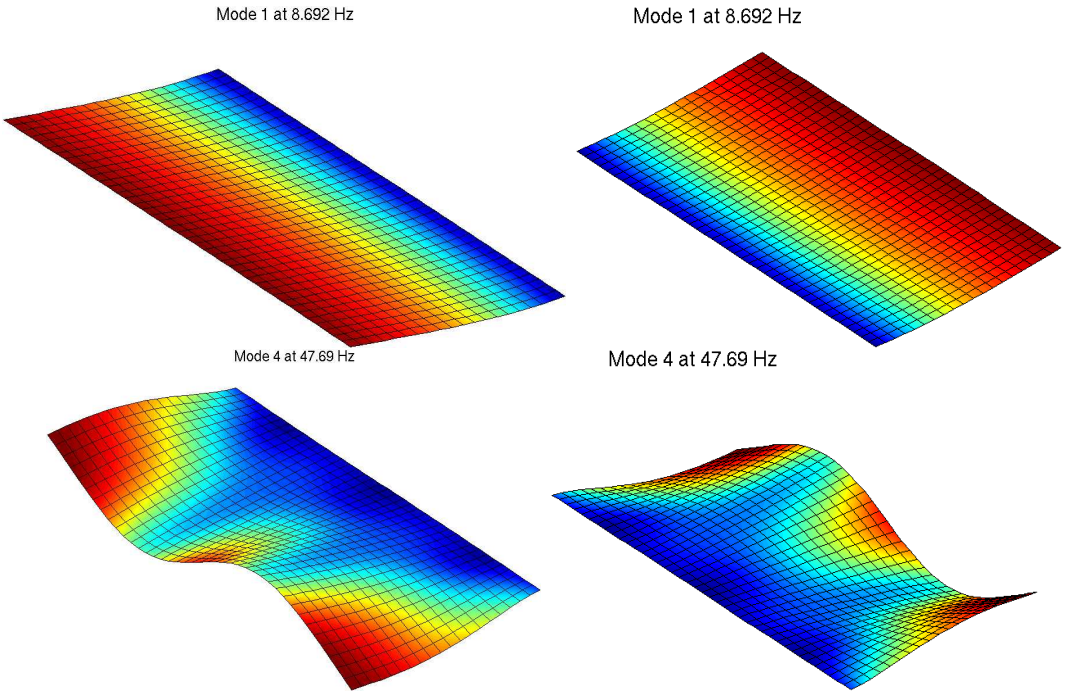
For applications in dynamics and static condensation, the assumption that loads are concentrated on the interface is no longer valid. Inertial effects from the complementary DOF of both components now need to be taken into account. Their representation is obtained by using a basis of Rayleigh-Ritz vectors for both q_{c1} and q_{c2} displacement respectively named $[T_{c1}]_{N_{1c}\times n_{1c}}$ and $[T_{c2}]_{N_{2c}\times n_{2c}}$

$$[T_R]_{(N_{1c}+N_i+N_{2c})\times(n_{1c}+N_i+n_{2c})} = \begin{bmatrix} T_{c1} & -K_{1cc}^{-1}K_{1ci} & 0 \\ 0 & \mathcal{I} & 0 \\ 0 & -K_{2cc}^{-1}K_{2ci} & T_{c2} \end{bmatrix} \tag{2.10}$$

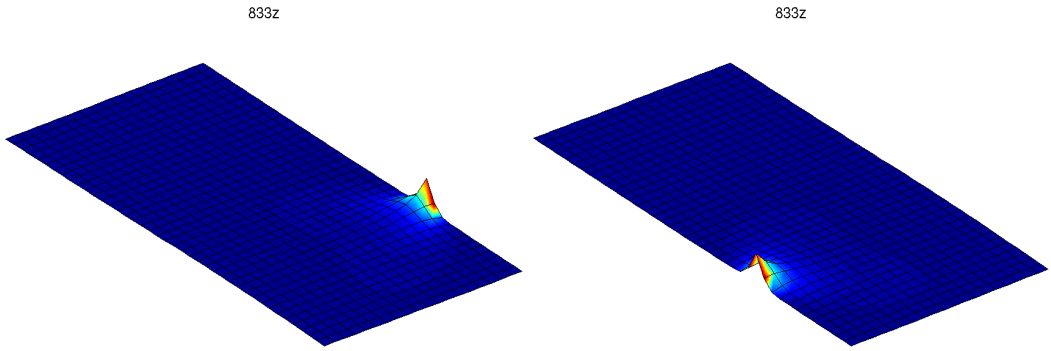
Usually, the reduction bases T_{c1} and T_{c2} are chosen so that the projected complementary DOF blocks of K_{cR} are diagonal. The natural reduction basis is then the normal modes ϕ_{j1} and ϕ_{j2} of the initial complementary DOF blocks $1cc$ and $2cc$. These blocks correspond the matrices K_1, M_1 and K_2, M_2 whose interface DOF have been removed, that is to say fixed. The relation

$$\begin{cases} T_{R1}^T Z_{1cc} T_{R1} = \left[\backslash\omega_{j1}^2\backslash\right] \\ T_{R2}^T Z_{2cc} T_{R2} = \left[\backslash\omega_{j2}^2\backslash\right] \end{cases} \tag{2.11}$$

consequently corresponds to the fixed interface modes of structures 1 and 2. The method is commonly known as the *Craig-Bampton* reduction. For the plate example, fixed interface modes are computed for both components, sampled in figure 2.3a. The constraint modes, or static responses to unit displacements enforced on the interface, are computed separately and shown in figure 2.3b.



(a) Fixed interface Modes



(b) Constraint modes (here z motion of node 833)

Figure 2.3: Sample fixed interface half-plate modes for the CMS and constraint modes

The coupled reduced matrices are then of size $N_R = (N_i + n_{1c} + n_{2c})$. For mass normalized fixed interface modes, the reduced stiffness matrix is then given by

$$[K_R]_{N_R \times N_R} = \begin{bmatrix} \begin{bmatrix} \backslash\omega_{j1}^2 \backslash \end{bmatrix} & 0 & 0 \\ 0 & K_{1ii} + K_{2ii} - K_{1ic}K_{1cc}^{-1}K_{1ci} - K_{2ic}K_{2cc}^{-1}K_{2ci} & 0 \\ 0 & 0 & \begin{bmatrix} \backslash\omega_{j2}^2 \backslash \end{bmatrix} \end{bmatrix} \tag{2.12}$$

and the mass by

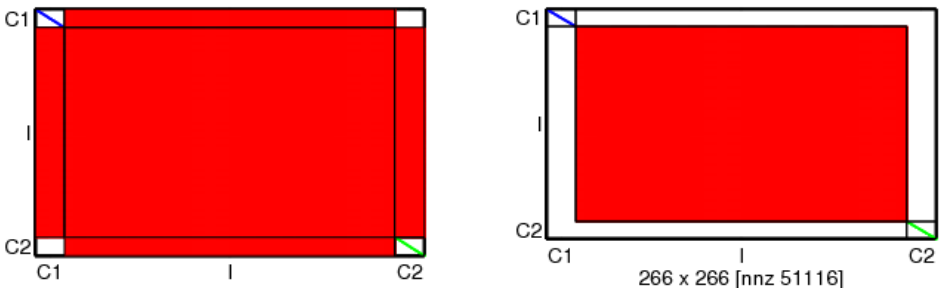
$$[M_R]_{N_R \times N_R} = \begin{bmatrix} \mathcal{I} & M_{1ci} - M_{1cc}K_{1cc}^{-1}K_{1ci} & 0 \\ M_{1ic} - K_{1ic}K_{1cc}^{-1}M_{1cc} & \hat{M}_{ii} & M_{2ic} - K_{2ic}K_{2cc}^{-1}M_{2cc} \\ 0 & M_{2ci} - M_{2cc}K_{2cc}^{-1}K_{2ci} & \mathcal{I} \end{bmatrix} \tag{2.13}$$

with

$$\begin{aligned} \hat{M}_{ii N_i \times N_i} = \; & M_{1ii} + K_{1ic}K_{1cc}^{-1}M_{1ci} + M_{1ic}K_{1cc}^{-1}K_{1ci} + K_{1ic}K_{1cc}^{-1}M_{1cc}K_{1cc}^{-1}K_{1ci} + \\ & M_{2ii} + K_{2ic}K_{2cc}^{-1}M_{2ci} + M_{2ic}K_{2cc}^{-1}K_{2ci} + K_{2ic}K_{2cc}^{-1}M_{2cc}K_{2cc}^{-1}K_{2ci} \end{aligned} \tag{2.14}$$

The reduced mass and stiffness matrix topologies given in (2.12) and (2.13) are presented in figure 2.4. They feature small diagonal blocks at the level of each component. The interface coupling blocks are however full. The mass matrix in particular bears full outer-diagonal blocks driven by the number of interface DOF and the number of modes kept in each component. In the example, 20 fixed interface modes are kept, the interface shows 246 DOF. The matrix size is reduced by 97%, from 10,086 DOF to 286 DOF. The number of non zero terms is less decreased. Results are good for the stiffness matrix showing a decrease factor of 83% from 307,212 to 51,116 – increasing the number of kept interface modes will not generate many more terms. The mass matrix shows a 60% decrease from 292,459 to 115,752, the final number of elements is however driven by the number of fixed interface modes, due to the non diagonal coupling terms.

To keep reasonable interface sizes, multilevel substructuring can be employed, which is the objective of the AMLS detailed in section 2.4. For applications in transient analysis this is however not acceptable and an alternative will be proposed in the next section.



(a) Reduced Mass, NZ=115,752 (b) Reduced Stiffness, NZ=51,116

Figure 2.4: Topology of the assembled mass and stiffness matrices using Craig-Bampton reduction for the plate split in two. 20 interface modes kept for both half-plates

The accuracy of the Craig-Bampton method is controlled through the selection of the number of retained constraint modes. For the plate example, figure 2.5 shows the reduced model convergence when increasing the number of retained constraint modes. The method very rapidly achieves reasonable accuracy of a few percent, but, as visible in figure 2.5 convergence to a higher accuracy is then slower requiring to keep fixed interface modes with frequencies much above the frequency band of interest. The figure thus shows that for a 1% tolerance, the number of modes well predicted increases sharply with the number of fixed interface modes kept with possibly more system modes well predicted than interface modes kept. For a tolerance of 0.1%, the number of modes to keep is however already much higher. In reality, good accuracy is achieved by keeping constraint modes within the target bandwidth. But keeping more modes only very slowly improves convergence.

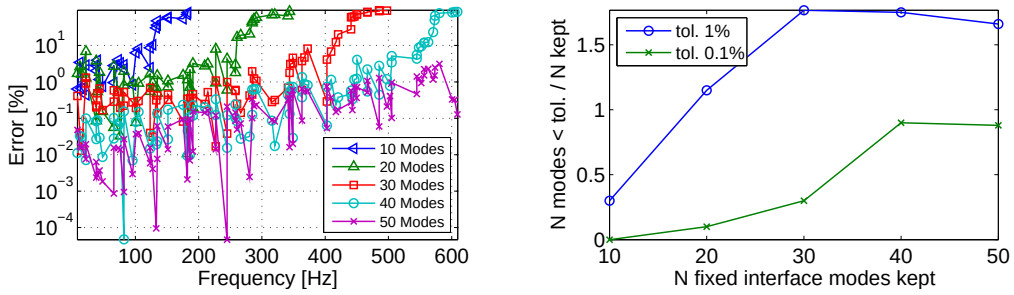


Figure 2.5: Convergence of the 2 components plate reduced by the Craig-Bampton method as function of the number of fixed interface modes kept for each component. Left: frequency error [%]. Right: fraction of modes under the tolerance against the number of interface modes kept

The good initial accuracy is sufficient for many applications and this has motivated a widespread use of the Craig-Bampton method. The following section will however show how full system modes can be used to generate reduced models that give exact modes and do not present large interface matrices.

2.2.2 Coupling by physical interface - exact reduced model

The size of interfaces is a major limitation of the Craig-Bampton method that will be discussed again for multi-level reduction in section 2.4. Considering disjoint components coupled by a physical interface is an alternative that can lead to much smaller models as will be shown here.

The second limitation addressed here is that accuracy of CMS is linked to its classical key assumption: the reduction occurs at the component level. While this assumption is necessary for the classical objective of reducing initial computations, the objective is here to reduce the computational cost of time integration in section 2.2.5, and run rapid design studies involving component properties in section 2.3. As in [20], one will thus seek a reduction that reproduces full system modes exactly.

The principle of the method is to consider disjoint components. Reduction bases are then written using reduction bases defined by block

$$\{q\}_{N_{1c}+N_{2c}\times 1} = [T_R]\,\{q_R\}_{n_{1c}+n_{2c}\times 1} = \begin{bmatrix} T_{c1} & 0 \\ 0 & T_{c2} \end{bmatrix} \begin{bmatrix} q_{R1} \\ q_{R2} \end{bmatrix} \tag{2.15}$$

and the dynamic stiffness can be decomposed as the sum of independent component contributions Z_i and an interface coupling matrix Z_I

$$[Z]_{(N_{1c}+N_{2c})\times(N_{1c}+N_{2c})} = [Z_{el}] + [Z_I] = \begin{bmatrix} Z_1 & 0 \\ 0 & Z_2 \end{bmatrix} + \begin{bmatrix} Z_{I11} & Z_{I12} \\ Z_{I21} & Z_{I22} \end{bmatrix} \tag{2.16}$$

In [20], physical coupling was achieved by automatically generating a layer of elements at the interface between components. A penalization approach is also possible but with the usual difficulties of selecting proper parameters.

In the applications of this work, the physical components are linked by contact law based interactions. The interface coupling thus corresponds to the stiffness generated by the linearization of the contact load around the static equilibrium and will be called an *interaction* matrix. This is detailed in section 3.2.4.

In this formulation, continuity obviously is always verified since (2.15) expresses motion using the full model FEM coordinates. The reduced dynamic stiffness matrix $[Z_R]$ is written

$$[Z_R]_{(n_{1c}+n_{2c})\times(n_{1c}+n_{2c})} = \begin{bmatrix} T_{c1}^T Z_1 T_{c1} & 0 \\ 0 & T_{c2}^T Z_2 T_{c2} \end{bmatrix} + \begin{bmatrix} T_{c1}^T Z_{I11} T_{c1} & T_{c1}^T Z_{I12} T_{c2} \\ T_{c2}^T Z_{I21} T_{c1} & T_{c2}^T Z_{I22} T_{c2} \end{bmatrix} \tag{2.17}$$

From the interaction representation of figure 2.1b, the elastic dynamic stiffness considers each components independently. Its topology is block diagonal and simply contains the dynamic stiffness of each component. The interaction matrix represents the component coupling associated to each interface DOF and in the present case only stiffness is used. Its topology is by block, a block i, j being non null if components i and j are interacting.

Unlike the Craig-Bampton method which needed complete interface description to enforce continuity, the method only needs to focus on selecting retained shapes relevant for other objectives. Reduction bases $[T_{ci}]_{N_i\times n_i}$ can be generated using the restriction of the system modes on each component, $\left[\Phi_{|ci}\right]$. In such case, the reduced system will give the exact result. The subspaces generated for each component are effectively coherent with the assembled system

subspace and all interface motion is well described since by definition all the desired motion is retained in the basis.

In this approach, the diagonal blocks of the elastic matrix $[Z_{el}]$ are not diagonal by default, since the system modes are orthogonal with respect to the assembled matrix $[Z]$. An orthogonalization procedure [20] is thus used to improve basis conditioning and obtain diagonal matrices.

2.2.3 Using component modes as DOF

To obtain a richer component description, the modes of the components themselves can be used. Thus with

$$[T_{ci}]_{N_{ic}\times n_{ic}} = [\phi_{ci}] \tag{2.18}$$

one has

$$\begin{cases} T_{c1}^T Z_1 T_{c1} = \phi_{j1}^T Z_1 \phi_{j1} = \left[\backslash \omega_{j1\backslash}^2\right] \\ \\ T_{c2}^T Z_2 T_{c2} = \phi_{j1}^T Z_1 \phi_{j1} = \left[\backslash \omega_{j2\backslash}^2\right] \end{cases} \tag{2.19}$$

and the coupled system equations

$$[Z_R]_{(n_{1c}+n_{2c})\times(n_{1c}+n_{2c})} = \begin{bmatrix} \left[\backslash \omega_{j1\backslash}^2\right] & 0 \\ 0 & \left[\backslash \omega_{j2\backslash}^2\right] \end{bmatrix} + \begin{bmatrix} T_{c1}^T Z_{I11} T_{c1} & T_{c1}^T Z_{I12} T_{c2} \\ T_{c2}^T Z_{I21} T_{c1} & T_{c2}^T Z_{I22} T_{c2} \end{bmatrix} \tag{2.20}$$

The elastic matrix of equation (2.20) is diagonal and features each component free/free modes square pulsations on the diagonal. This feature is of great interest as these values now become explicitly present at the assembly level. The objective of reducing design cycles between local component parameters and assembly behavior estimations thus becomes possible.

The reduction (2.15) assumes a displacement subspace for each interface, which will not contain all possible displacement. In an extreme case, one would need to use as many Rayleigh-Ritz vectors as there are DOF on the interface and lose the method interest. The problem is then to find a subspace diagonalizing component matrices while proposing an interface displacement subspace representing all possible movements from interactions with other components.

The space of the full system real modes Φ provides a subspace containing all relevant information for the interface movements. From a single component point of view, the basis of real free/free modes is complemented by the trace of the real system modes on the component, noted $\Phi_{|ci}$ as proposed in section 2.2.2. Computing the full modes of the assembled system at a nominal state is deemed accessible (through AMLS at least) and computationally acceptable.

The additional information introduced mainly concerns the component interaction with others and thus provide an adequate subspace containing the component free/free modes and all relevant interface movements. For component i, the reduction basis is written

$$[T_{ci}]_{N_{ic}\times n_{ic}} = \left[\begin{array}{cc} \phi_{ci} & \Phi_{|ci} \end{array}\right]_{Orth.} \tag{2.21}$$

To avoid redundant data, equation (2.21) specifies an orthogonalization at the component level with respect to mass and stiffness, so that

$$\begin{cases} [T_{ci}]^T [K_i] [T_{ci}] = \left[\backslash \omega_{i,j\backslash}^2\right] \\ \\ [T_{ci}]^T [M_i] [T_{ci}] = \mathcal{I} \end{cases} \tag{2.22}$$

Since the orthogonalization is performed at the component scale, the component free/free modes verify (2.22) by definition. Only the contribution of the enrichment modes is made orthogonal to the component modes.

The topology of the assembly matrices is then optimal, with the system mass matrix being the identity and the elastic stiffness being diagonal with the square of free/free component pulsations. The interaction matrix only shows full blocks for interacting components.

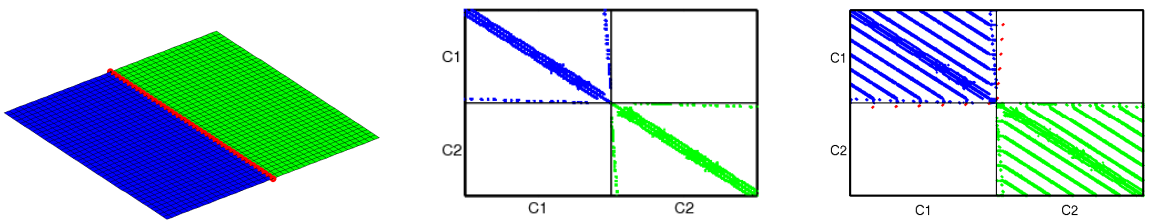
The first remarkable feature is that the reduced system thus produced features exactly the same modes than the full one – to numerical precision. Indeed, the assembled real modes information has been used explicitly and can therefore be recovered. The second interesting feature is that modified components, that can be well represented by the nominal components modes, can be well predicted by the same reduced model. This provides the basis for the reanalysis procedures that will be illustrated in the following sections.

The main drawback is that, while the interaction matrix still features a block topology, non-null reduced blocks are however full. The generation of full blocks is inevitable using Rayleigh-Ritz vectors, but the problem is here much smaller than with the Craig-Bampton approach since the interfaces sizes are associated with the number of system and component modes and thus unrelated to the number of DOF in the interface topology.

2.2.4 A coupled plate example

The CMT method is illustrated using the plate shown in figure 2.2a. Although physical coupling is normally preferred, coupling of the two half plate models with interface DOF duplicated is obtained with a penalization stiffness of $10^7 N/m$ or $10^7 N/rad$ on all DOF. This value was chosen to provide equivalent full plate modeshapes and frequencies while not altering the system conditioning.

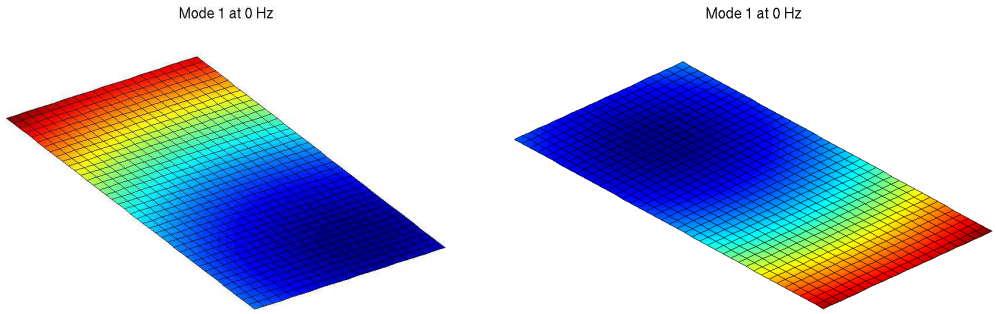
The plate, presented in figure 2.6 can then be assembled following the topology given by equation (2.16). Figure 2.6 also shows the elastic and coupling stiffness. It must be noticed that the coupling matrix diagonal blocks are typically not empty.



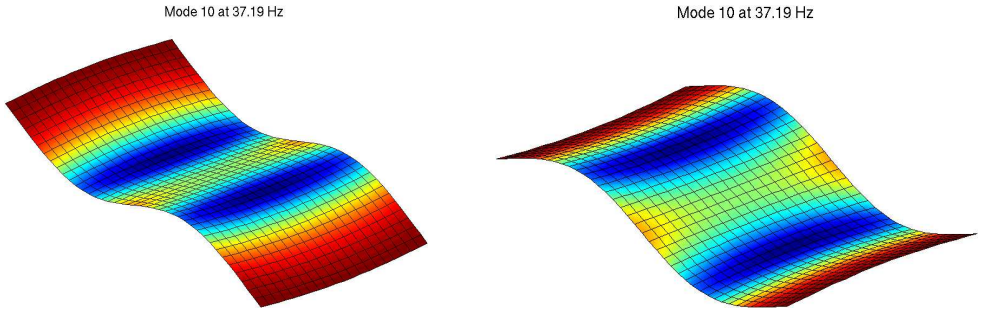
(a) Plate model (b) K_{el}, NZ=596,370 (c) K_{int}, NZ=309,425

Figure 2.6: Finite element matrix topologies for the CMT formulation. Half-plate #1 in blue, #2 in green. Interface made of spring elements in red

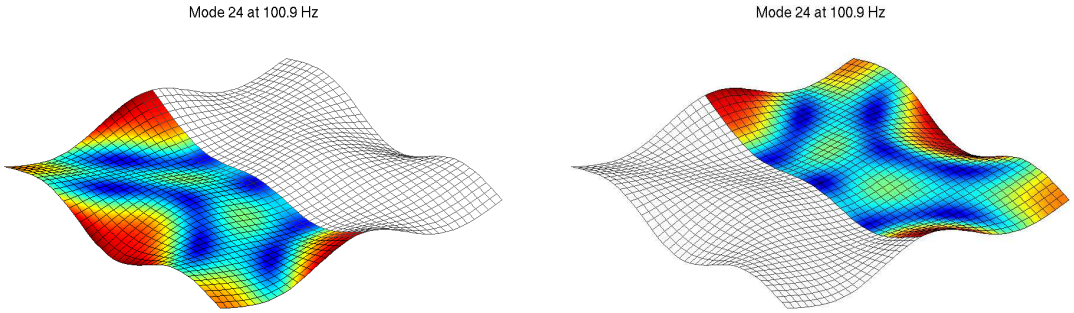
Using assembled system modes and component modes for both half-plates in free/free conditions, the enhanced reduction basis presented in equation (2.21) can be computed. Sample shapes are shown in figure 2.7. The free/free modes are first present, and contain rigid body modes, figure 2.7a and 2.7b.



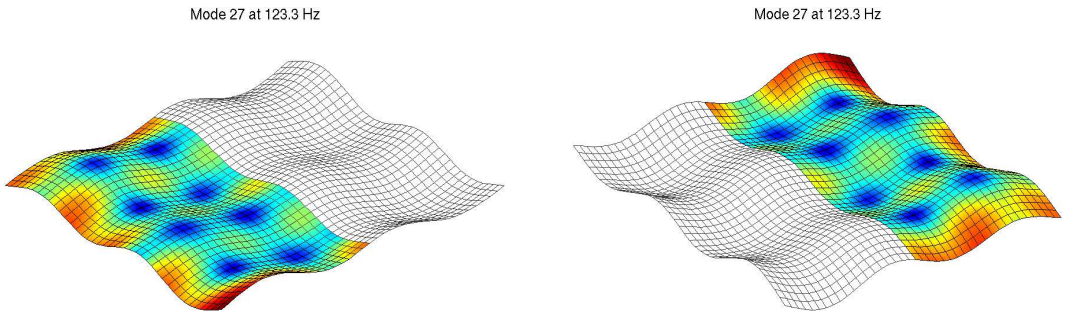
(a) Component free/free modes – #1



(b) Component free/free modes – #10



(c) Real mode trace #24



(d) Real mode trace #27

Figure 2.7: Sample half-plate free/free modes (including rigid body ones) and shape of the enhancement modes (trace of the assembly real modes). 20 component modes are kept for the example

The enhancement modes shapes are not exactly the trace of the assembled mode shapes plotted in figure 2.7c and 2.7d due to the basis orthogonalization. They typically contain most of their information at the interface. Enhancement shape #27, representing for the example the 7^{th} enhancement mode is shown in figure 2.8. The shape is clearly neither close to a free/free component mode nor to an assembled mode.

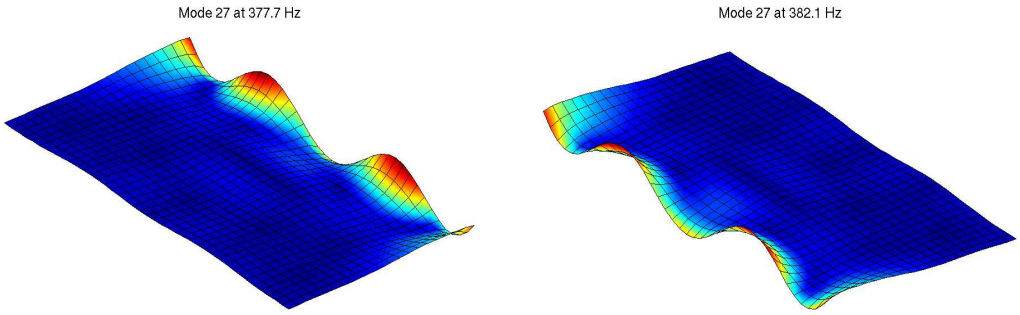
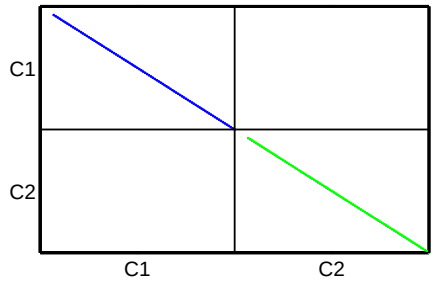


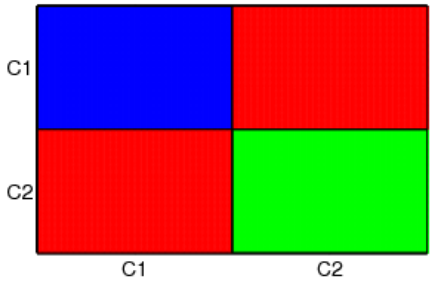
Figure 2.8: Sample enhancement shape of the orthogonalized basis

Frequencies can be computed for each enhancement mode due to the orthonormalization. Indeed, the same conditions are applied to the enhancement shapes than to the component free/free normal modes, so that an equivalent frequency coming from the values of the projected stiffness matrix diagonal can be exploited to illustrate the effect of the enhancement.

In the following, the reduction is performed using 100 assembled modes for enhancement and by using a variable number of component modes retained in the system, from 10 to 100. Due to the orthogonalization process and evident redundancy occurrences between the component free/free modes and assembled modes (*e.g.* rigid body modes), the number of modes kept in the final reduction basis is not the exact sum of all modes retained. The matrices presented in figure 2.6 can be projected and feature the specific topologies aforementioned.



(a) K_{elR}, NZ = 188



(b) K_{intR}, NZ = 40,000

Figure 2.9: Reduced matrix topologies for the CMT formulation. The use of rigid body modes yields zero diagonal terms on the elastic matrix. The mass matrix is reduced to the identity

The resulting topologies are very sparse, indeed the mass and elastic stiffness matrices are diagonal, the size of the number of modes kept, 200. Null diagonal terms appear in the elastic matrix for component rigid body modes. The interaction matrix is potentially full – interaction between every component is quite rare for real applications – but only driven by the number of modes kept for each component.

Looking at the diagonal of the reduced elastic matrix, the effect of the enhancement can be observed. Figure 2.10 clearly shows that the frequency increase of the enhanced modes is higher than the increase observed for the free/free modes themselves.

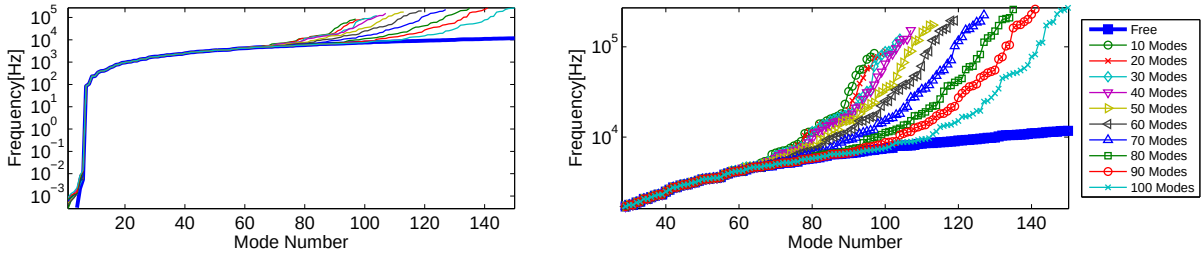


Figure 2.10: Diagonal of the elastic stiffness matrix block $(1,1)$ as function of the number of free/free modes kept. The free blue line corresponds to the component #1 free/free mode frequencies, as a reference. Left: global picture. Right: zoom in.

First, using the assembled system real modes allows recovering the component modes. In figure 2.10, the frequency of the enhanced modes match perfectly the component frequency until the 60^{th} mode shape.

For the frequency range of interest, the information contained in the component modes seems relevant until the 60^{th} shape, afterwards, the enhancement allows a quicker convergence. Using more than 60 component modes indeed generates enhanced modes of higher frequencies, as observed in figure 2.10, showing that the information added by more component modes is not the most pertinent. This is confirmed in figure 2.11, where the number of enhancement modes needed is compared to the number of component modes kept. The dotted lines are iso values for which adding one component modes reduces by one the number of enhancement modes.

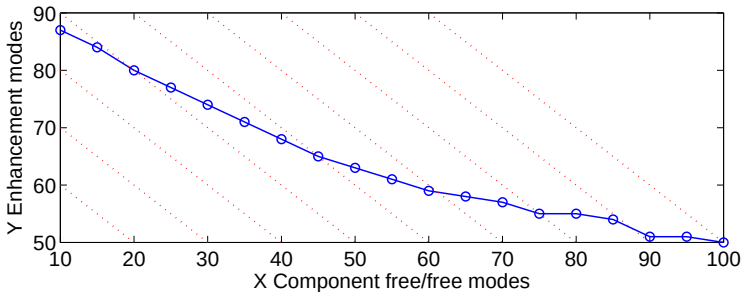


Figure 2.11: Number of enhancement modes kept by the orthogonalization process from the 100 retained as function of the number of component modes retained.

Clearly, the number of enhancement modes needed first decreases linearly as the information added by more component modes is increased, then their number tends toward a horizontal asymptote around 50 modes. This shows that a component free/free mode subspace of reasonable size alone is not sufficient to represent the interface displacement. Obviously the convergence using only components free/free modes will then be rather slow. Looking at an enhancement shape from figure 2.8 it is clear that such shape will be of very high frequency at the scale of the component.

To illustrate the fact that the lack of information from the component free/free modes subspace is not easily filled by increasing their numbers. Figure 2.12 shows the convergence results without using the enhancement modes. This method can be referred to the methods developed by MacNeal [49] or Rubin [50]. It can be seen first of all that the overall precision is not satisfying, and does not actually increase as function of the number of modes kept. The error of several percent in average is not kept for a large frequency band.

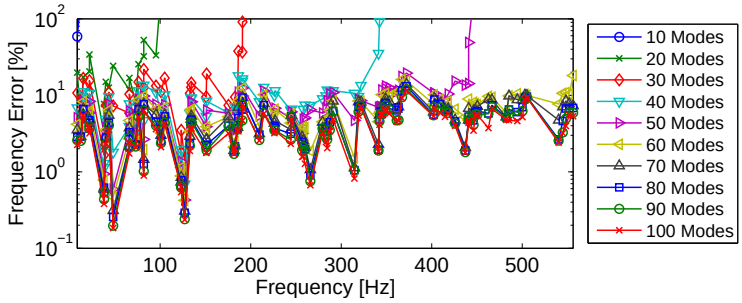


Figure 2.12: Frequency error using a reduction basis only featuring component free/free modes

Using the enhanced reduction basis yields an exact reduced system from the real mode point of view. Figure 2.13 indeed shows a precision over 10^{-7} for the whole frequency band – this has to be put in perspective with the maximal precision attainable of 10^{-8} for real eigenvalue solvers.

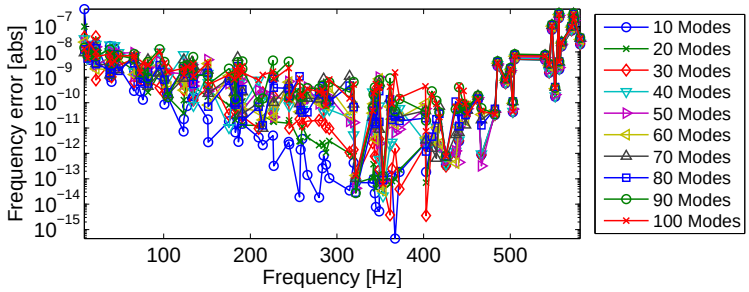


Figure 2.13: Frequency error between the full model and the reduced one, using the enhanced reduction basis

2.2.5 Efficient reduction for a system with local non-linearities

The concept of reduced models with exact real modes and implicitly reduced interfaces as presented in section 2.2.2 opens new ways for reduction strategies. In particular, the time simulation of systems with local non-linearities is one of the main goals of the thesis, which requires a model reduction method with different needs than a traditional CMS.

The reduced model is of course expected to be small, but also reasonably sparse and very precise on the non-linearity – for a brake system: the pad-disc interaction. The reduction will then couple a superelement with a finite element part where the local non-linearity is operating. The CMT application is then very relevant as the system at its nominal state will have the same modes than the full one, and the interface size between the superelement and the finite element part will not impact the system final size.

The illustration in this section will feature the final brake assembly presented in section 1.2 for clarity. The split between the superelement and the part kept non-linear is presented in figure 2.14. In addition to keeping the *non-linear* DOF taken as the DOF where non-linear contact-friction forces occur, the pad lining is also kept, with a section of the disc surface. Keeping the lining is based on the fact that the material stiffness difference with the disc may generate non-linear effects or wave propagations in the lining thickness.

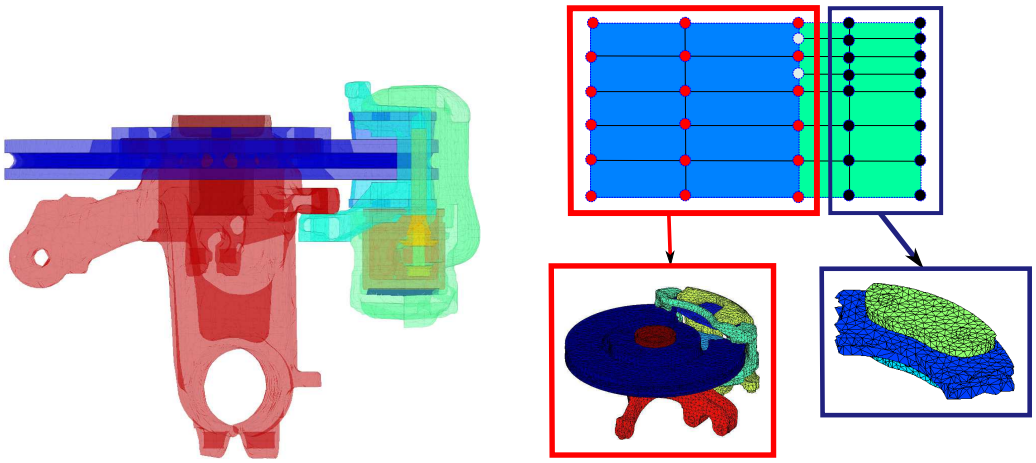


Figure 2.14: Left: View of the industrial brake model provided by Bosch. Right: Industrial brake model reduction strategy. The blue part is kept unreduced, the red part is reduced, also implicitly reducing the interface (red and green nodes)

Adapting the reduction method, two DOF sets are here distinguished, a *linear* part (q_b) that will be reduced (q_r), and a finite element part (q_c) that will not be altered. No component modes are necessary, the use of the system real modes at the static state are however kept for the reduction basis, written

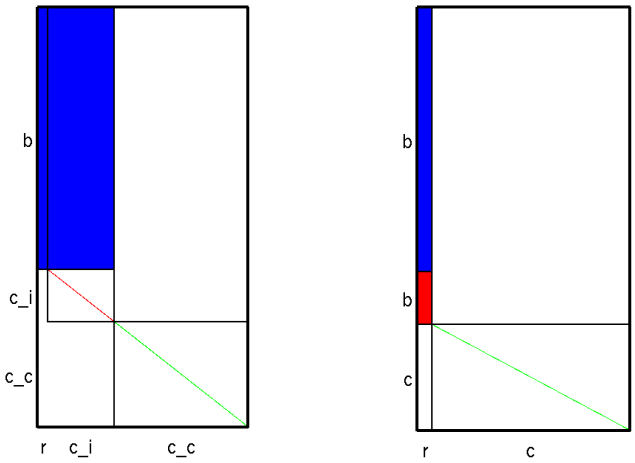
$$\left\{ \begin{array}{c} q_b \\ q_c \end{array} \right\} = \left[\begin{array}{cc} \left[[\Phi_{1:NM}]_{|b} \;\; q_{|0b} \right] & 0 \\ 0 & [I]_c \end{array} \right] \left\{ \begin{array}{c} q_r \\ q_c \end{array} \right\} \tag{2.23}$$

A Craig-Bampton reduction would proceed differently. The DOF retained would be divided in two sets, (q_{cc}) corresponding to the internal DOF chosen to be retained explicitly, and (q_{ci}) for the parts of retained elements connected to the reduced part. The reduction basis writes then

$$\left\{ \begin{array}{c} q_b \\ q_{c_i} \\ q_{c_c} \end{array} \right\} = \left[\begin{array}{ccc} [\phi_{Fixed}] & -K_{bb}^{-1} K_{bc_i} & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{array} \right] \left\{ \begin{array}{c} q_r \\ q_{c_i} \\ q_{c_c} \end{array} \right\} \tag{2.24}$$

In the present case, using a frequency cutoff at $20kHz$ to comply with the squeal frequency range, 250 system modes are to be retained for the superelement reduction. The reduction bases

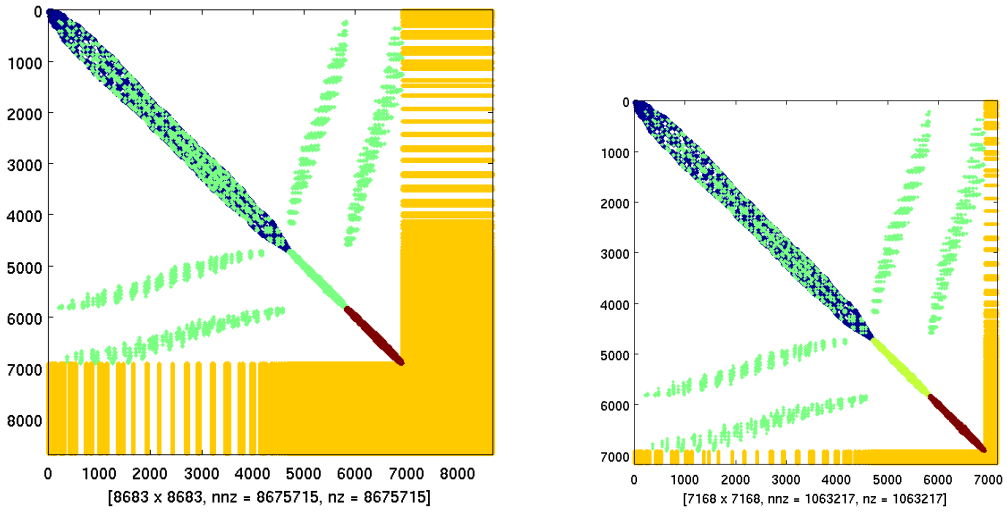
topologies from the new reduction equation (2.23) and the classical reduction equation (2.24) are presented in figure 2.15.



(a) Craig Bampton (b) New reduction

Figure 2.15: Reduction bases topology plots

The illustration clearly shows that a very large full block of the size of the interface is thus avoided when using (2.23). The matrix topologies yielded by both projection methods are shown in figure 2.16. Due to the model sizes, this illustration is limited to a model using 4 nodes linear tetrahedrons instead of the 10 nodes quadratic ones from the original mesh.



(a) Craig-Bampton (b) New

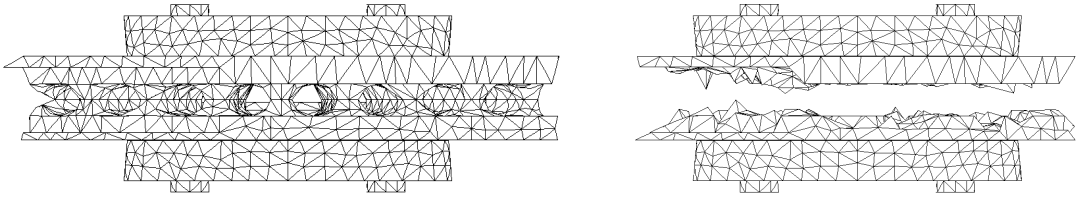
Figure 2.16: Reduced matrix topologies for the 4 nodes linear tetrahedron model. Left: Craig-Bampton reduction. Right: proposed reduction

The DOF associated to each component are distinguishable in figure 2.16. From left to right appear the disc section, outer pad lining, inner pad lining and reduced part. Outer diagonal terms correspond to contact coupling between the pads and disc.

Using reduction basis (2.23) yields a small superelement block, of size 251. The Craig-Bampton basis (2.24) appends to the 251 modes the whole interface, which is of size 1,765 (12,400 for the original 10 nodes quadratic tetrahedron model). The proposed method gives then a great advantage in terms of memory, as the new matrix is much sparser. The number of non null elements is indeed reduced by a factor eight. In application to the industrial brake

model, the memory size taken by the superelement block thus reduced from $1,000MB$ to $0.5MB$.

Further sparsity optimization has been carried out by considering the way of cutting the disc section. The pad lining material is much softer than the disc, and the contact has low effect on the disc behavior. Figure 2.17 plots the initial kept area keeping a complete disc section and the new reduction considering only the disc elements underlying the contact surface.



(a) 61,900 DOF (b) 37,157 DOF

Figure 2.17: Definition strategies for the non reduced area

Cutting the disc under the contact surface further increases the interface size, explaining the results presented in table 2.1 showing each reduction statistics. The matrix sizes are plotted against computing times for basic time integration operations (section 4.2). The mechanical residue computation requires a matrix vector product, while the solve operation requires the inversion to compute an increment correction.

Method	N_DOF	NNZ	SE_CONN	Residue	Solve
Disc full section	61,900	10,557,313	12,379	232ms	220ms
Elements under contact	37,157	13,058,917	21,508	204ms	162ms
Nodes under contact	30,768	9,484,043	15,219	150ms	127ms

Table 2.1: Models resulting from the reduction method, number of DOF, number of non zeros, number of DOF connected to the superelement for the reduced stiffness matrix, residue and solve times

Considering the elements underlying contact reduced significantly the model size but also increased the superelement connectivity, which in fact decreased the system sparsity. A better solution is to consider the reduction to the nodes underlying the disc contact surface, thus decreasing the final model size, but more importantly reducing the connectivity, so that the final sparsity is improved. Residue computation was thus reduced by 35% and solve times by 42%.

2.3 The Component Mode Tuning method (CMT)

Reduced models using component modes explicitly as DOF were introduced in section 2.2.3. A major motivation of this choice is to allow rapid design studies, where properties of a component are modified and predictions of the system level response are analyzed.

Unlike traditional CMS methods from the late 70's, the motivation is not to allow coupled analyzes of larger models. The AMLS algorithms, presented in section 2.4, are a modern evolution of these motivations and make the computation of nominal responses for very large models acceptable. Considering this evolution, one thus considers that computing the exact response of a nominal structure is a completely acceptable step. The reduction proposed in 2.2.3 uses these results to generate a reduced model that is exact for the nominal parameters. It consequently allows optimizing design cycles.

The proposed design methodology is typical of *Structural Dynamics Modification* (SDM) methods. One seeks alterations of the system mass and stiffness properties that improve the

system level vibration performance. Testing modifications using direct methods, where full analyses are used for each configuration, is unrealistic with computation times of a few hours for each configuration. The proposed reduced model allows sensitivity studies and reanalysis iterations within minutes.

The suggested reduction method provides a system model that contains component modes explicitly. Tuning component properties is thus possible trough sensitivity and reanalysis, hence the name of Component Mode Tuning (CMT).

The combination of exact nominal values and very fast sensitivity computations and reanalysis is the key of the proposed approach that will be illustrated using a plate example, section 2.3.1 and an industrial brake model, section 2.3.2.

2.3.1 Sample CMT uses – plate example

As a first illustration of the CMT, one will consider modifications that impact the damping characteristics of the system complex modes. When the system has unstable modes, redesign is performed to modify the modal behavior in terms of frequency or damping. Classification of research directions can be achieved by evaluating the sensitivities $\frac{\partial \lambda_j(p)}{\partial p}$ of the system eigenvalues to a parameter p, here taken as the frequency of a component mode.

From appendix A, the complex mode sensitivity expressed by equation (A.13) can be simplified. For the variation of the frequency of a component mode,

$$\begin{cases} \frac{\partial [M]}{\partial p} = 0 \\ \frac{\partial [C]}{\partial p} = 0 \\ \frac{\partial [K]}{\partial p} = diag(\langle 0 \cdots 0 \ 1 \ 0 \cdots 0 \rangle) \end{cases} \tag{2.25}$$

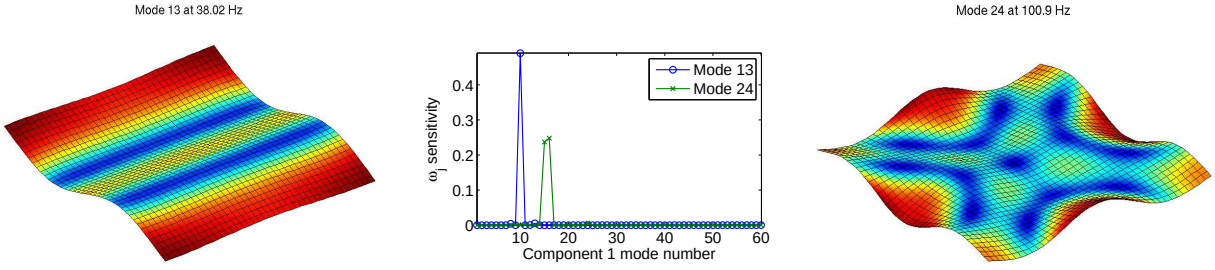
where the stiffness variation has a single non-null value standing for the component mode to vary. The sensitivity computation becomes then

$$\frac{\partial \lambda_j}{\partial p} = \left\{ \psi_{lj}^T \right\} [k_p] \{ \psi_{rj} \} \tag{2.26}$$

where $[k_p] = diag(\langle 0 \cdots 0 \ 1 \ 0 \cdots 0 \rangle)$, and ψ_{lj} (resp. ψ_{rj}) is the j^{th} left (resp. right) system complex mode.

Equation (2.26) is used to sort component modes (diagonal terms on k_p) by their impact on the complex eigenvalues of the assembled system. Since the sensitivity is not constant, reanalysis – evaluation at non-nominal design points – will need to be addressed later.

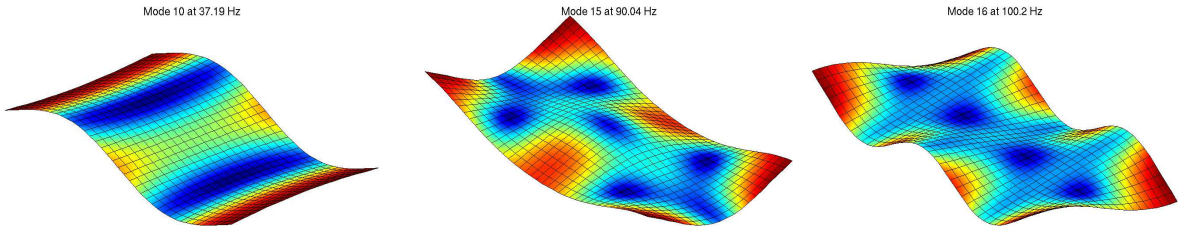
The sensitivity of real mode frequencies of the full plate modes is computed for all half-plate #1 component modes in figure 2.18.



(a) Plate mode #13 (b) Sensitivities (c) Plate mode #24

Figure 2.18: Sample sensitivity computation for plate modes #13 and #24. 60 component modes are retained for the reduced model and 100 assembled complex modes

It can be seen that sample plate mode #13, figure 2.18a is only sensitive to component mode #10, plotted in figure 2.19a. Similarly, mode #24 in figure 2.18c is sensitive to two component modes, #15, figure 2.19b and #16, figure 2.19c. Two component modes from the first half-plate at 90 and 100 Hz are here coupled to yield mode #24 at 101 Hz.



(a) Semi plate mode #10 (b) Semi plate mode #15 (c) Semi plate mode #16

Figure 2.19: Sensitive component modes for plate modes #13 and #24: half-plate modes #10, #15 and #16.

These results are rather simple due to the plate symmetry properties but illustrate the mechanisms of CMT. Besides, the method relevance is highlighted by showing that the component mode DOF are the ones responding in the study, and not the assembled enhancement ones.

The second illustration of the CMT is focused on reanalysis. One first considers a Young modulus variation of the half-plate #1, second, the effect a local mass addition on a node of half-plate #1.

The Young modulus variation is very easy to test as the system presented features an isotropic material. The component matrix is thus proportional to the Young modulus. That is

$$K_{var}(E) = K_{elvar}(E) + K_{int} = \begin{bmatrix} (E/E_0)\,[K_1(E_0)] & 0 \\ 0 & K_2 \end{bmatrix} + \begin{bmatrix} K_{I11} & K_{I12} \\ K_{I21} & K_{I22} \end{bmatrix} \tag{2.27}$$

Applying equation (2.27) to the system reduced at its nominal state ($E = E_0$), the frequency variation due to the Young modulus variation can be studied quickly, and with enough design points to track each mode evolution, as presented in figure 2.20.

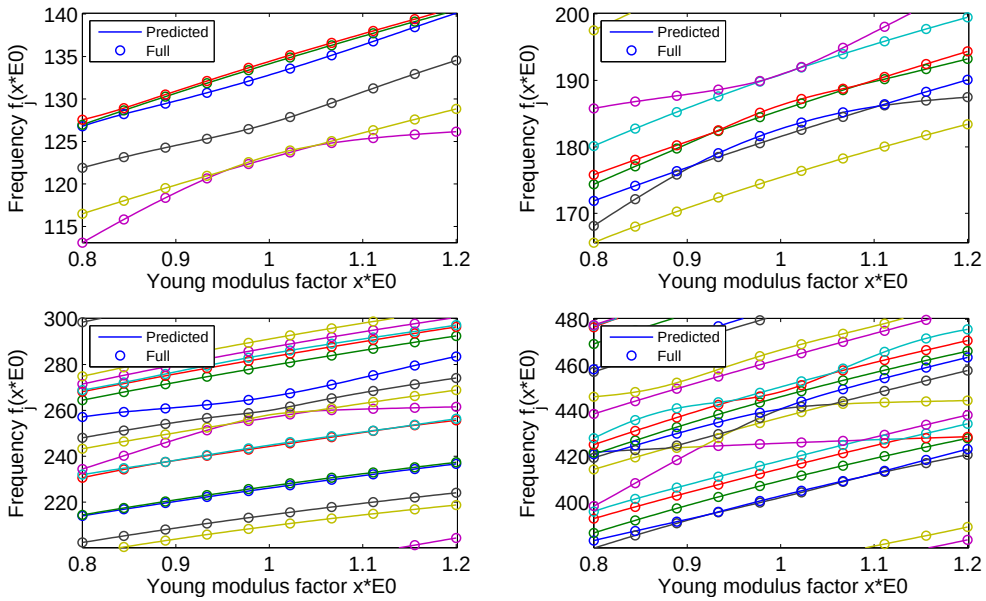


Figure 2.20: Plate frequency variation as function of the Young modulus variation of half-plate #1, predictions from the CMT (-) and recomputations (o)

For a rather significant Young modulus variation, from -20% to +20% the predictions by the CMT are very precise. The fully recomputed points coincide very well with predictions. Figure 2.21 shows in particular that for a 20% increase of Young modulus, the prediction error is of maximum 0.07% for a frequency variation due to the modification around 10%. Reanalysis allow studies of mode crossing and veering patterns at a very low computational cost.

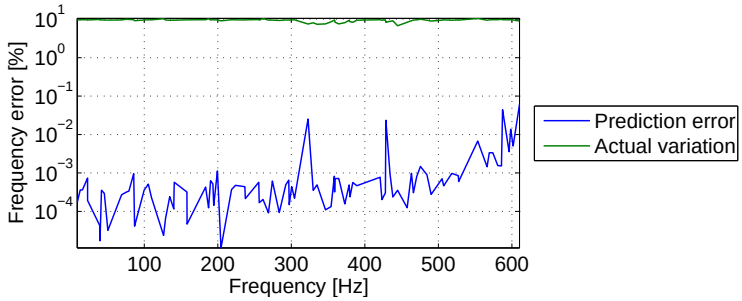


Figure 2.21: Prediction frequency error against recomputed frequency shift, for the second half-plate Young modulus increase by 20%

Figure 2.22 illustrates the effect of the +20% Young modulus variation on the modeshapes, which can be dramatically altered. Indeed, the plate assembly symmetry along the splitting axis is broken (modes computed in free/free conditions). The shapes presented can even be very different from one side of the plate to another.

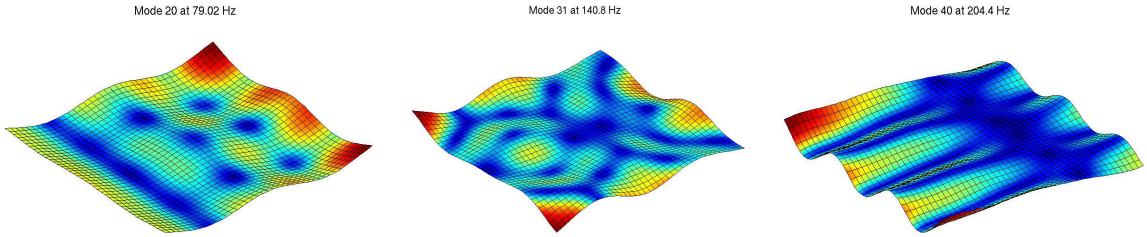


Figure 2.22: Plate mode shape variation for an increase by 20% of half-plate #1 Young modulus

The mass modification is performed by adding a local mass to a node of component half-plate #1 (noted m_1), as presented in figure 2.23. The local mass added varies from 0 to 0.1 grams, from an initial nodal mass of 0.018 grams. A modified mass matrix is thus computed,

$$M_{var}(m_1) = M + \Delta M = M + \begin{bmatrix} T_{c1}^T \Delta M_1 T_{c1} & 0 \\ 0 & M_2 \end{bmatrix} \tag{2.28}$$

where ΔM_1 represents the mass addition m_1 in the component.

Figure 2.23 illustrates the loss of symmetry along the component coupling axis due to the additional mass.

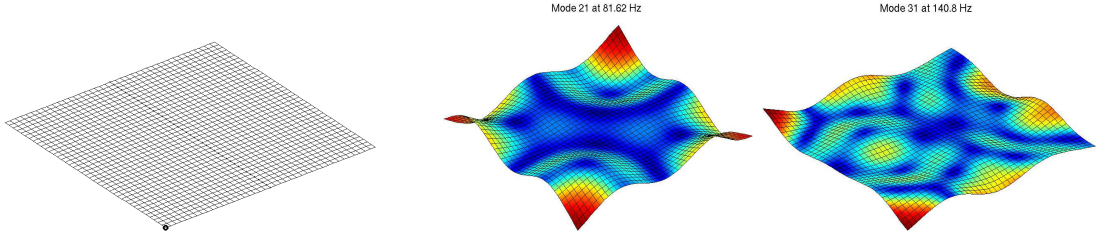


Figure 2.23: Plate node bearing the mass modification. Sample affected mode shapes plotted for the recomputed modes with an additional mass of 0.1 gram.

The effect on selected frequencies, prediction and recomputation, is presented in figure 2.24. Most modes are insensitive (less than 0.1% of variation). The exact error for the 0.1 gram increase is shown in figure 2.25 with the actual frequency shift due to the mass addition. The variation, quite extreme but very local can be almost null for certain modes, and very significant for others.

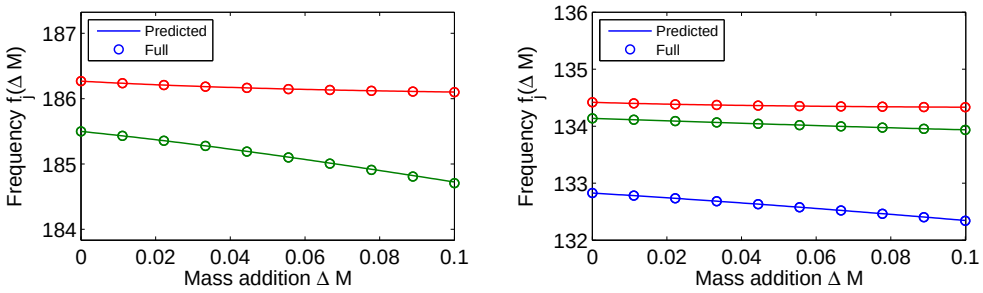


Figure 2.24: Plate frequency variation as function of the added mass on half-plate #2, predictions from the CMT (-) and recomputations (o)

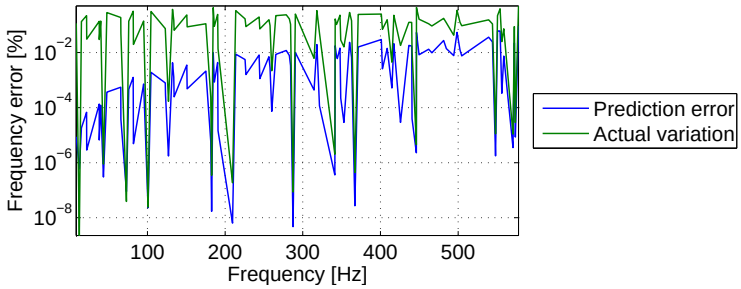
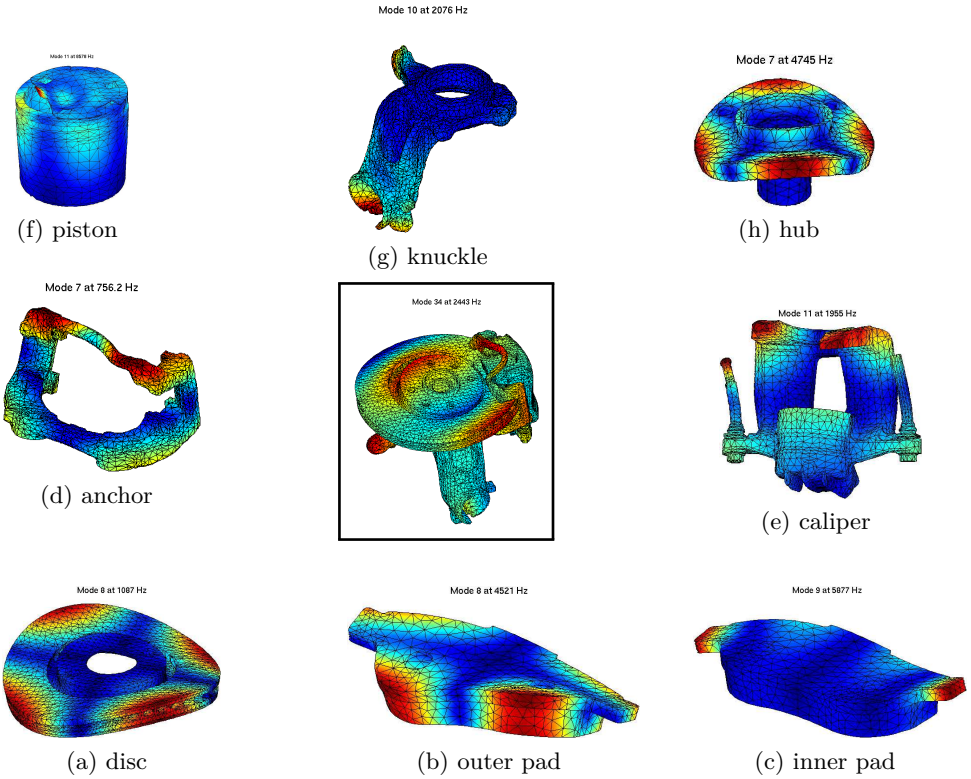


Figure 2.25: Prediction frequency error against recomputed frequency shift for a 0.1 gram increase on the half-plate #2 node #1

The prediction error is the smallest for the largest frequency shifts actually observed. The fact that the prediction error is in general a few orders below the actual frequency shift proves the prediction relevance. Modes unaltered by the mass variation show negligible frequency shifts, which make the error evaluation difficult, they are nevertheless detected by the method.

2.3.2 An industrial brake application

The CMT method is now applied to the case of the industrial brake model provided by Bosch. The model is presented in section 1.2, and sample component modes are illustrated in figure 2.26. Validations are performed for three specific studies, and highlight the method relevance for real life cases.



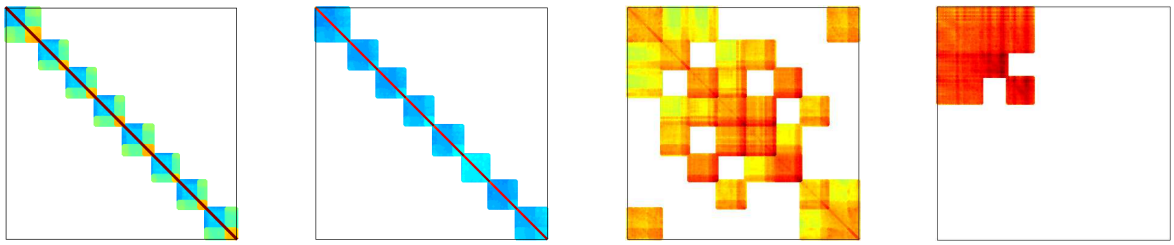
(f) piston (g) knuckle (h) hub

(d) anchor (e) caliper

(a) disc (b) outer pad (c) inner pad

Figure 2.26: Real Brake assembly and components. Modeshape illustrations

The reduction procedure, described in details in section 2.2.3, is applied. After computing the system real modes, each component is treated separately and transformed into a superelement, before a final assembly. The matrix topologies of the reduced model are presented in figure 2.27. The 8 component are clearly seen in the first two matrices, where out of diagonal terms are negligible. The interaction is split in a symmetric part and an unsymmetric part that only involves the pad/disc interaction (see section 3.2.4 for contact-friction formulation details).



(a) Mass matrix (b) Elastic stiffness matrix (c) Symmetric interaction (d) Unsymmetric coupling

Figure 2.27: Real brake reduced matrices. Logarithmic color scale from low values in blue to largest values in dark red

Depending on the component stiffness the number of kept modes varies. The final system size is of 1,300 DOF instead of the initial 600,000 DOF.

The real modes of the reduced model are expected to be the same than the ones of the full model since they are explicitly input in the reduction basis. Figure 2.28 shows that, despite expected round-off errors, precision is under 10^{-5} and below 10^{-8} for much of the frequency range. No comparison is possible for the rigid body modes since their frequency is theoretically zero.

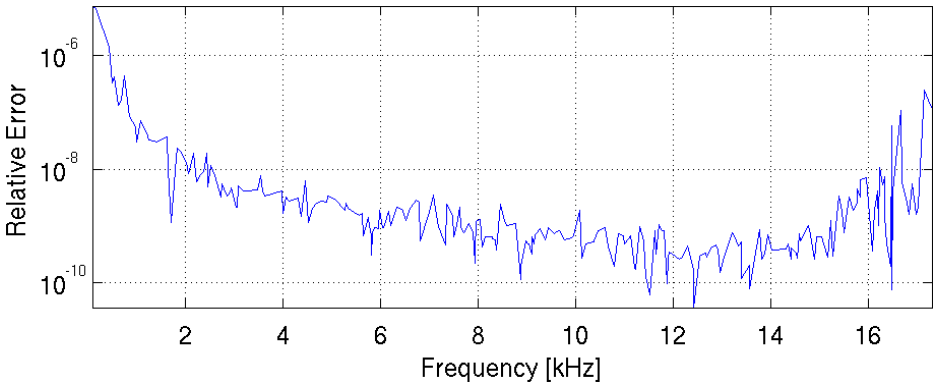


Figure 2.28: Difference between the full brake model and the reduced model: real frequencies

A major interest of the CMT is its ability to predict non nominal responses with good accuracy. This will be illustrated here with a strong modification to the nominal model. A Young modulus variation of the disc is first tested, for real and complex mode predictions. Shifting the Young modulus of the disc material implies to modify the corresponding block of the elastic stiffness matrix. In our case, block $(1,1)$ of the matrix figure 2.27b is given by equation (2.27). The real modes predicted are compared to the real modes of the modified model recomputed on the full model.

Figure 2.29 shows that the relative frequency error between the full modified model and the prediction is under 2%. In most cases it is smaller than the relative frequency shift observed. It can be noticed that the higher the relative shift, the smaller the prediction error. SDT and ABAQUS results are presented separately as a verification of the SDT solver, which Bosch wanted to validate.

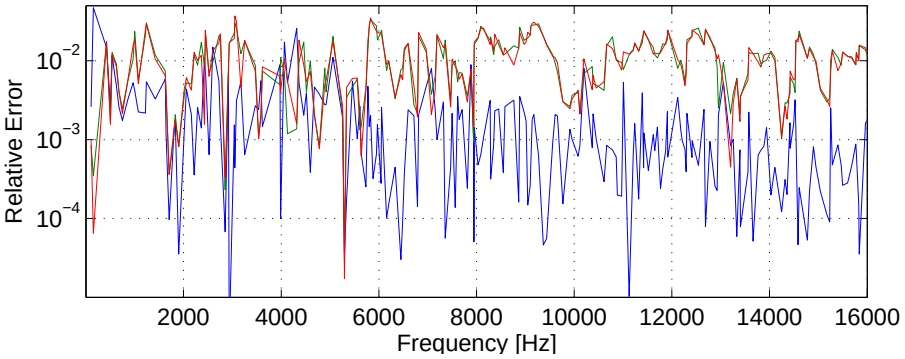


Figure 2.29: Relative error on the real modes predicted by the reduced model and a full ABAQUS recomputation for a 10% increase of the disc Young modulus. [–] Relative error on the frequencies. [–] Relative frequency shift computed with SDT. [–] Relative frequency shift computed with ABAQUS

Figure 2.30 validates that the frequency shift prediction error is smaller than the actual frequency shift observed for the modification. First $\Delta f_{abq} = \|f_{nom} - f_{abq-mod}\|$ is computed for the nominal ABAQUS model and the model recomputed using the disc Young modulus modification. Δf_{SDT} is computed using the frequency of the nominal model (proved to be identical to the nominal ABAQUS model) and the prediction in SDT. The shift error is then computed with $\Delta(\Delta f) = \|\frac{\Delta f_{abq} - \Delta f_{SDT}}{\Delta f_{abq}}\|$.

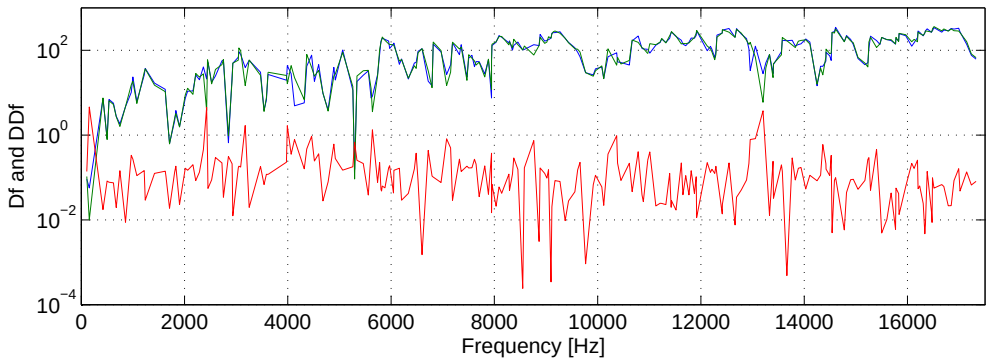


Figure 2.30: Frequency shift error. [–] Shift between the ABAQUS models (Δf_{abq}). [–] Shift between the SDT models (Δf_{SDT}). [–] Relative error between the predicted shift and the actual shift computed with ABAQUS ($\Delta(\Delta f)$)

After validating real mode predictions, one now seeks to verify the accuracy on complex modes by plotting a stability diagram as function of the disc Young modulus for a variation to nominal from -20% to +30%. The evolution is compared to a full recomputation in ABAQUS with a modified disc material with Young modulus shifts of -20%, +10% and +20%. The results of the prediction using the reduced model are compared to the ABAQUS results.

The complex mode stability diagrams are close although not overlaying due to the approximate solutions computed in both cases. Figure 2.31 shows details of the stability diagram. The trends are well followed and the stable/unstable transitions are properly detected.

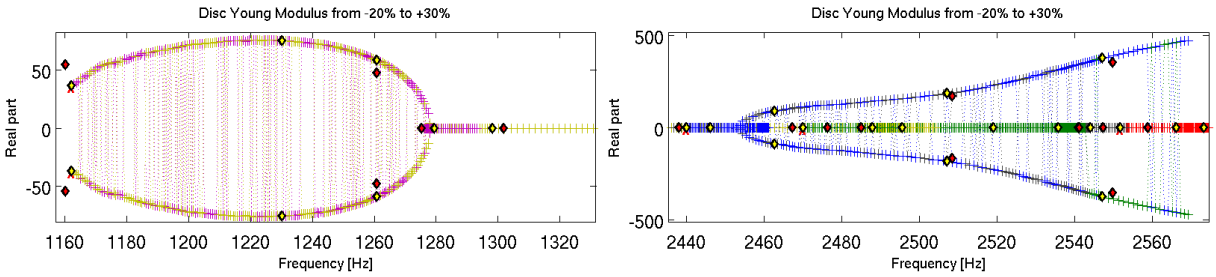


Figure 2.31: Full stability diagram as function of the disc Young modulus, green diamond at the nominal point, red diamonds for the ABAQUS recomputed solution and yellow diamonds for the corresponding prediction

As a second validation example, mass variations will be considered since they are common practice in SDM methods. A local mass variation is then tested on a specific component. The brake anchor component is suggested by Bosch. The effect of mass variations on this component is easy to observe.

The mass modification is conducted using equation (2.28). A local mass is added to the anchor at its *handle*. In practice, a point mass of 9 g is applied to node 400438 shown in figure 2.32. The effect on the non rigid anchor modes is shown in figure 2.33. Depending on the mode, a variation from -6% to nothing on the real frequencies is observed.

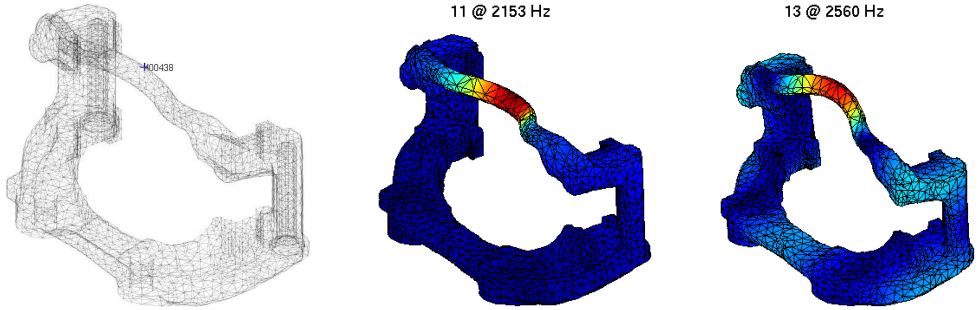


Figure 2.32: Local mass modification of the Anchor handle. Added mass location and the first two altered modes

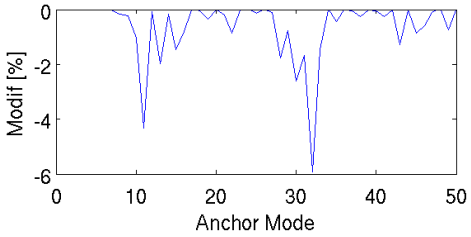


Figure 2.33: Effect of a mass modification on the Anchor on its real frequencies

The effect on the assembled modes is of course mainly associated to movements of the anchor handle. The prediction is accurate as shown in figure 2.34, where a 0.1% minimum precision is achieved after the rigid body modes. Figure 2.35 shows that the frequency shift prediction error is lower than the frequency shifts of the ABAQUS or SDT computation.

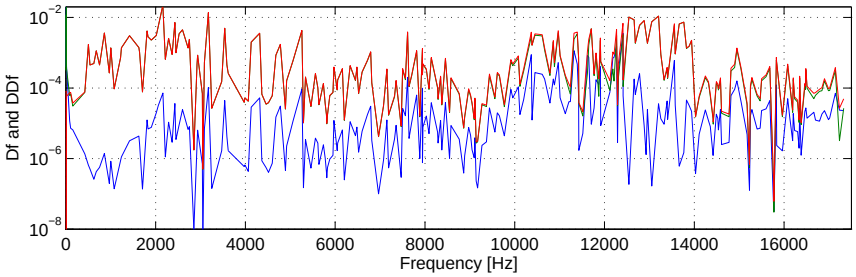


Figure 2.34: Relative error between the 250 first real modes of the ABAQUS recomputed solution and the SDT prediction. [–] Relative error on the frequencies. [–] Relative frequency shift computed with SDT. [–] Related frequency shift computed with ABAQUS

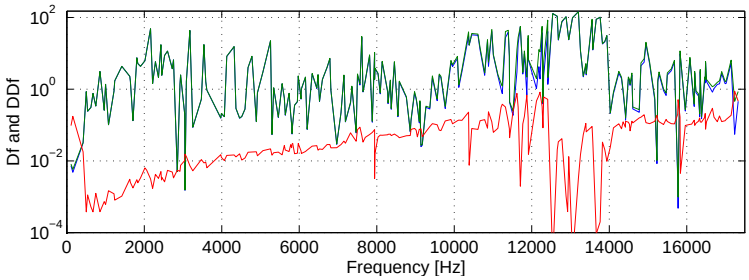


Figure 2.35: Frequency shift error. [–] Shift between the ABAQUS models (Δf_{abq}). [–] Shift between the SDT models (Δf_{SDT}). [–] Relative error between the predicted shift and the actual shift computed with ABAQUS ($\Delta(\Delta f)$)

The transverse Young's modulus of the lining is another classical design parameter used by Bosch for model updating. The stiffness matrix is not proportional to this coefficient, as was the case for the isotropic Young's modulus. In a classical approach [42], one uses a linear interpolation between two material settings to study the effect of parameter variations.

In application to the component reduction for brake models, two versions of the pad stiffness matrix are computed, then reduced using the same reduction basis. The choice of a nominal point for the reduction basis must be made. Computations have been performed for the lowest and highest moduli. Both points can be taken as the nominal point.

The lining transverse modulus used were from the static $E_{zz} = 275\ MPa$ to the dynamic modulus set at $E_{zz} = 3000\ MPa$. The results using both strategies are presented in figure 2.36. They are globally very close, but as expected, differences are visible locally, which is shown in the zoom-ins of figure 2.37.

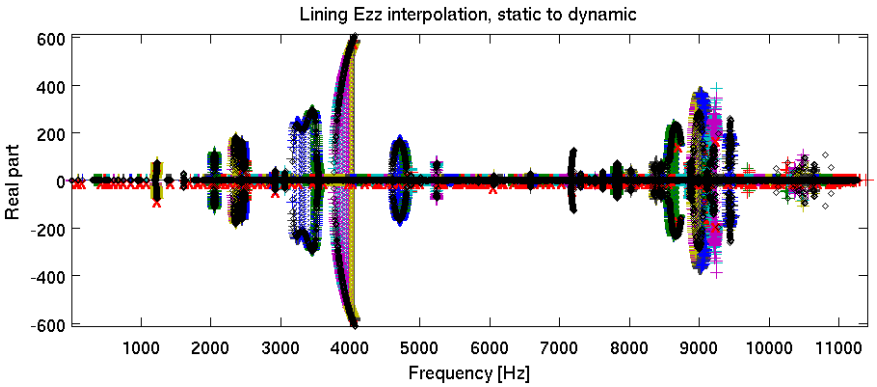


Figure 2.36: Stability diagram as function of the lining transverse Young modulus. (+) makers stand for the high E_{zz} value, ($\diamond$) for the low value used in the reduction basis

Figure 2.37 shows that differences are more visible for higher moduli, corresponding to the highest frequency points of each mode. The results of both methods for lower moduli are closer, stiffer conditions should then be taken as a nominal point.

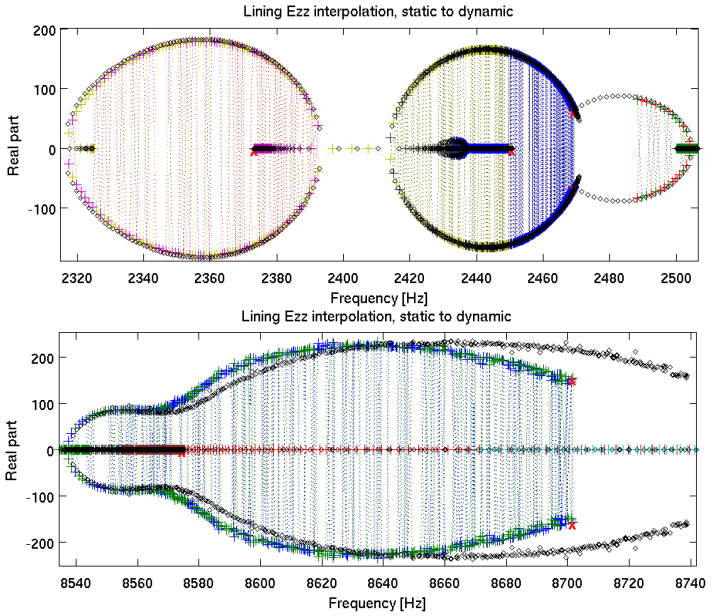


Figure 2.37: Details of the stability diagram as function of E_{zz}. (+) makers stand for the high E_{zz} value, ($\diamond$) for the low value used in the reduction basis

Possible improvements to this study would enhance the nominal reduction basis by combining the assembled modes of the second extreme configurations. A gain in precision is expected when using such method. Appendix C presents studies with enhanced reduction bases for friction and braking pressure parametric studies.

The CMT method shows here a real interest in industrial applications. Indeed, very precise predictions can be obtained from one single nominal state and the free/free modal bases of each component. As the effect of a local component modification can be quickly evaluated at the assembly level, the cost of a design cycle can be dramatically reduced. This eventually opens new perspectives for assembly updating in correlation to experimental measurements.

2.4 AMLS concept: iterative coupling of substructures

One of the original motivations for CMS methods was to allow computation of larger models. The last decade has seen a renewed interest in this objective using automated multi-level substructuring methods initiated in the work of Kaplan *et al.* [28]. Implementations for the use of such concepts are available for frequency response function and real mode computations in the major commercial codes (NASTRAN, ABAQUS, ANSYS, ...) but with few associated publications.

As the computation of system modes is a required input of the proposed reduction methods, an AMLS type algorithm was thus implemented. Implementation issues have limited the performance achieved, so the present section only seeks to analyze the process and key issues that need to be addressed. Section 2.4.1 addresses partitioning methods that are used to automatically generate a hierarchic division in components. Section 2.4.2 details the base equations for AMLS. Implementation issues are eventually discussed in section 2.4.3.

2.4.1 Automatic generation of interfaces

Classical linear system resolution methods require the computation of an LDU decomposition. Computation costs become however quickly prohibitive for large problems. The main idea is then to partition the problem in a series of smaller and much less costly uncoupled problems, combined with a coupling problem whose cost needs to be controlled. Partitioning is thus the generation of interfaces between subdomains.

To partition a finite element model, the connectivities of the system must be studied. This can be achieved either by working on the finite element mesh or on the system stiffness matrix. Working on the matrix is preferable as the target application for automatic partitioning concerns applications like the AMLS. The term *solver* implies a low level tool whose input is clearly defined and independent from the model. The direct use of matrices is therefore preferred.

The system partitioning is performed thanks to a front elimination tree. Such method is classically used to factor matrices since it optimizes storage space, operations and subsequent fill-in patterns. Front elimination trees are based on graph theory, they can therefore be computed in our case for mesh topologies or matrices.

Several techniques exist as presented for example in [51, 52]. They are based on graphs associated with the stiffness matrix for which vertices are the DOF and edges correspond to non zero coupling terms between two DOF. The procedure then seeks a way of minimizing connections and of finding a single root.

Three canonical ways of graph partitioning exists, briefly detailed in [51], one based on level set ordering techniques (Cuthill-McKee algorithm), which basically ranges partitions as a leveled neighborhood from an initial vertex. Independent set ordering techniques are an improvement with an increased robustness as neighborhood levels are set from all vertices. These techniques are the base of the minimum degree algorithm which was widely used in the last decade.

Computer performance improvement helped the emergence of other techniques based on graph dissection, and in particular bisection. Spectral bisection is based on the evaluation of the eigenvalues of a matrix of identical topology containing -1 at non diagonal non zero values and the vertex degrees on diagonal terms. The Fiedler vector (the eigenvector corresponding to the second eigenvalue) contains in roughly equal parts negative and positive values which allows a bisection of the graph. this procedure can then be computed iteratively on the subparts, yielding the Recursive Spectral Bisection method. This method is known to produce high quality partitions.

Less costly bisection methods are based on graph traversing. The graph is traversed from an initial vertex until half the vertices have been seen, which is a direct graph partition. The use of pseudo peripheral nodes allows starting from low degree (with low connectivity) vertices which are prone to give partitions of equal sizes. Again the method is applied iteratively on the subgraphs obtained. This method weakness in terms of graph quality and computation costs is covered by the use of a multi level approach at the graph level. Using a coarsened graph to begin computations and refining in subgraphs afterwards proved to be very efficient and is the base technique of the METIS library [53].

The SPOOLES library [54] is robustly linked to SDT and provides similar methods. Its most interesting implementation concerns the use of a hybrid approach exploiting either nested dissection or multisection. Multisection is a generalization of bisection, which will either start with an initial multisector as the root, or perform local multisections from a bisected graph.

Partitioning can be well illustrated in application to a finite element plate model. Figure 2.38 shows a 10 by 10 plate mesh split into 4. The model matrices can then be reordered by blocks to follow the tree ordering, yielding specific matrix topologies

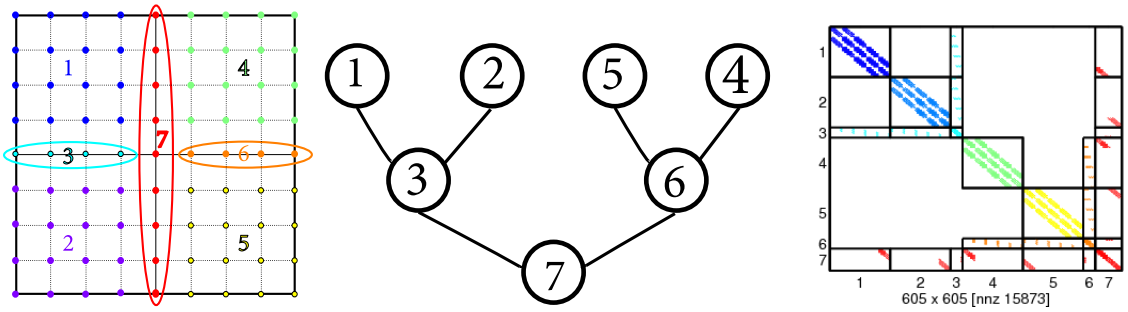
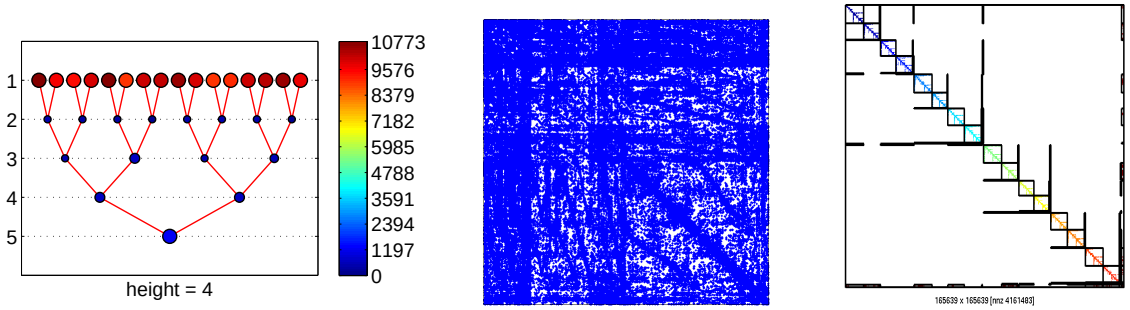


Figure 2.38: Plate sample partitioning, the corresponding elimination tree and reordered matrix

The result of matrix partitioning is represented under the form of an elimination tree. Each of the tree's vertices stand for a substructure, *i.e.* a subset of DOF of the original model. The tree vertices can be sorted in two categories, *leaf* vertices and *separator* vertices. The tree leaves are substructures without any child and can be compared to usual CMS substructures (1, 2, 4, 5 in the example). The intermediate leaves (3, 6) and the root (7) (top level, no parent vertices) are separators, and play the role of the leaves interfaces. The separators are considered as actual substructures with proper DOF sets.

The reordered matrix topology follows an *imbricate arrow* shape. Coupling blocks can be non zero only if a child/parent dependence path exists between the two. No *a priori* assumption can be made about the density of the coupling blocks, however it is likely that only a small subset of the child DOF is connected to the parent DOF. Besides, a block is not necessarily directly connected to every one of its parents (or ancestors).

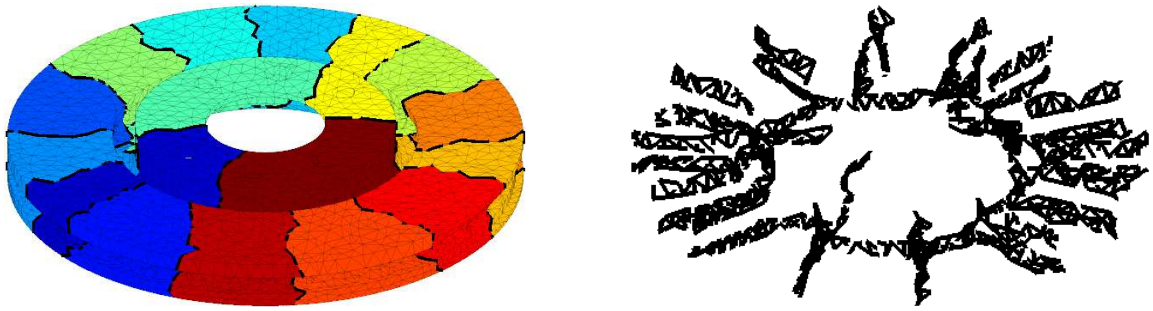
As an industrial application, a disk model is used. The disc component is the largest of the brake assembly and has 165,639 DOF. The stiffness matrix density is of $4.4.10^{-2}$%, and the initial mass matrix topology is presented in figure 2.39b.



(a) Front Elimination tree (b) Initial topology (c) Front topology

Figure 2.39: Disc mass matrix initial and reordered topologies with elimination tree. Tree color code: vertex DOF size

The disc shows a good connectivity repartition due to its relative cyclic symmetry, hence the well balanced binary elimination tree obtained. The corresponding substructure topologies are plotted in figure 2.40, in which leaves are shown as colored elements and separators as bold black lines.



(a) Colored leaves (b) Separators

Figure 2.40: Disc partition from the elimination tree of figure 2.39a

2.4.2 Formulation of an iterative CMS

The goal of CMS in the present application is to find a basis in which the stiffness matrix K is block diagonal, using a classical LDL^T decomposition. All coupling terms then go to the mass matrix M which is used for the eigenvalue solution.

Due to the reordering performed by one of the methods presented in section 2.4.1, K and M are block partitioned, written as

$$K = [K_{ij}]_{1 \leq i,j \leq N} \tag{2.29}$$

A Crout factorization of K is thus sought to obtain a block diagonal stiffness matrix D, (see [51] for example), such that

$$K = LDL^T \tag{2.30}$$

where D is block diagonal, and L is a lower block triangular matrix. L is then used to project the mass matrix into $\hat{M}$. L and D terms can be identified from the general form of the LDL^T product, writing

$$K_{ij} = \left[LDL^T\right]_{ij} = \sum_{0<k\leq \min(i,j)} L_{ik} D_{kk} L^T_{jk} \tag{2.31}$$

and

$$M_{ij} = \left[L\hat{M}L^T\right]_{ij} = \sum_{k\leq i} L_{ik} \sum_{p\leq j} \hat{M}_{kp} L^T_{jp} \tag{2.32}$$

From the specific matrices topology many terms can be simplified in (2.31) and (2.32). In particular, the conditions on intermediary indices can be written as functions of Parent and Children groups, respectively written for a node i, $\mathcal{P}_i$ and $\mathcal{C}_i$. These notations follow the tree representations from figure 2.38.

A block by block iterative identification from leaf to root is then possible for L and D, yielding

$$L_{ij} = \begin{cases} \hat{K}_{ji} D_{ii}^{-1} & \text{if} \quad j \in \mathcal{P}_i \\ \mathcal{I} & \text{if} \quad i = j \\ 0 & \text{otherwise} \end{cases} \tag{2.33}$$

where for $i \neq j$ and thanks to the base matrices topology

$$\hat{K}_{ij} = K_{ij} - \sum_{k \in \mathcal{C}_i} \hat{K}_{ik} L_{jk}^T \tag{2.34}$$

and

$$D_{ii} = K_{ii} - \sum_{j \in \mathcal{C}_i} L_{ij} D_{jj} L_{ij}^T \tag{2.35}$$

It can be observed that intermediate stiffness coupling elements $\hat{K}$ appear in the process, although simplified in the resulting D matrix. Mass matrix terms can eventually be obtained by projection or also iteratively identified as

$$\hat{M}_{ij} = M_{ij} - \sum_{k \in \mathcal{C}_i} \sum_{p \in \mathcal{C}_j \cup \{j\}} L_{ik} \hat{M}_{kp} L_{jp}^T - \sum_{k \in \mathcal{C}_j} L_{ii} \hat{M}_{ik} L_{jk}^T \tag{2.36}$$

Equations (2.33)-(2.36) allows the computation of all terms of the projected matrices. The possibility of reducing all tree nodes can then be discussed independently.

In the classical SDT notations corresponding to the CMS nomenclature and in coherence with the rest of this chapter, one identifies the L matrix to a recursive projection matrix $\tilde{T}$, for which

$$\tilde{T}_{ij} = \begin{cases} -D_{ii}^{-1} \hat{K}_{ij} & \text{if} \quad j \in \mathcal{P}_i \\ \mathcal{I} & \text{if} \quad i = j \\ 0 & \text{otherwise} \end{cases} \tag{2.37}$$

In the following, matrices and vectors superscripted with a tilde denote recursively written bases. The recursive notion is kept for the presentation as it is the base of the algorithms classically presented in the literature.

The recursive pattern can simply be illustrated by the fact that coupling occurs between non consecutive levels. The plate example given in figure 2.38, shows in particular that each leaf is connected to its parent by one side and to the root by the other side. The restitution is then formulated in a recursive way, such that to be fully restituted the tree vertex needs all its parents to be already restituted, which can be written

$$q_i = \tilde{q}_i + \sum_{j \in \mathcal{P}_i} \tilde{T}_{ij} q_j \tag{2.38}$$

It follows a recursive definition of the restituted projection basis

$$T_{ij} = \tilde{T}_{ij} + \sum_{k \in \mathcal{P}_j} \tilde{T}_{ik} T_{kj} \tag{2.39}$$

The plate example provided in section 2 helps as an illustration. To recover the full displacement on leaf vertex #1, the restitution must already have been proceeded for the root vertex #7 and the intermediate separator #3. The leaf displacement depends on the displacement of

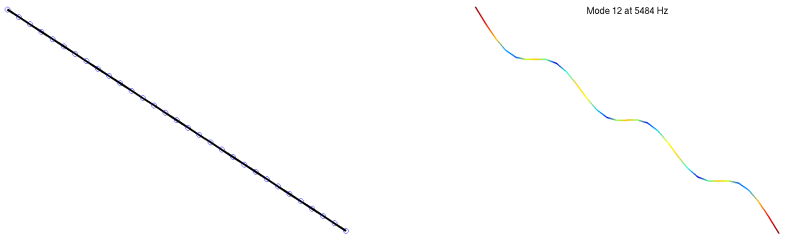
#3 but also of #7. Recursively, the displacement of #3 depends on #7 as well, which generates cross coupling terms,

$$\begin{cases} q_7 = & \tilde{q}_7 \\ q_3 = & \tilde{q}_3 + \tilde{T}_{37} q_7 \\ q_1 = & \tilde{q}_1 + \tilde{T}_{13} q_3 + \tilde{T}_{17} q_7 \end{cases} \tag{2.40}$$

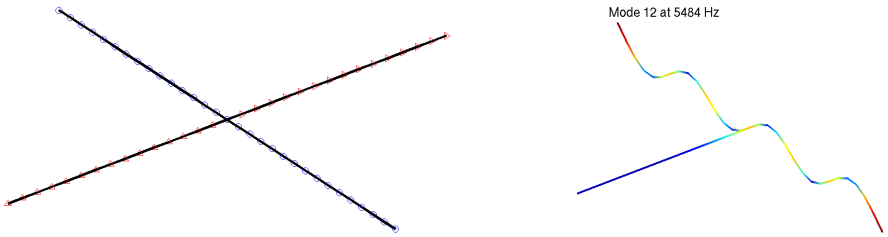
As a result the full restitution on q_1 is

$$q_1 = \tilde{q}_1 + \tilde{T}_{13}\tilde{q}_3 + \left(\tilde{T}_{17} + \tilde{T}_{13}\tilde{T}_{37}\right)\tilde{q}_7 \tag{2.41}$$

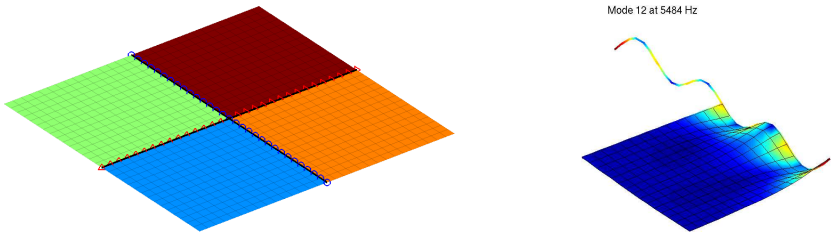
which is plotted in figure 2.41.



(a) Root level result



(b) Intermediate level restitution



(c) Leaf level restitution for #1

Figure 2.41: Sample restitution of a displacement of the root to its first and second level children

At this stage, an iterative Guyan condensation has been realized for a structure decomposed into sub-domains. The final condensed matrix is therefore exact as no reduction has been attempted and the modes can be computed on the root condensed matrices. Although theoretically interesting, such approach is underperforming as the condensation operations quickly become intensive. From the leaf upper level, all non null block matrices become full.

The point of view adopted here is of the multi-level reduction substructures. One seeks then a reduction basis diagonalizing the stiffness matrix, whose topology follows the initial stiffness topology.

$$[T] = [T_{ij}]_{1 \leq i,j \leq N} \tag{2.42}$$

where $[T]_{ij} \neq [0]$ if $j \in \mathcal{P}_i$ and $j > i$. The basic diagonalization is obtained as detailed above. Classical reduction procedures can then be exploited to diagonalize each block on its modes. Using equation (2.39) to represent the full condensation basis, the system reduction writes

$$[T_R] = [T]\,[\Phi] \tag{2.43}$$

where $[\Phi]$ represents the modes ϕ_i of each vertex on which all lower level condensations have been performed, which can be represented as

$$[T_R] = \begin{bmatrix} \mathcal{I} & \cdots & T_{1i} & \cdots & T_{1N} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & \mathcal{I} & \cdots & T_{iN} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & \mathcal{I} \end{bmatrix} \begin{bmatrix} \phi_1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & \phi_i & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & \phi_N \end{bmatrix} \tag{2.44}$$

The matrices obtained by the reduction phase are much smaller in size than the initial ones. Besides, the reduced stiffness matrix $\mathcal{K}$ can now be diagonal, with

$$\begin{cases} \mathcal{K}_{ii} = \phi_i^T D_{ii} \phi_i = \begin{bmatrix} \backslash w_i^2 \backslash \end{bmatrix} \\ \mathcal{K}_{ij} = 0 \qquad\qquad\qquad\qquad i \neq j \end{cases} \tag{2.45}$$

The mass matrix diagonal blocks are diagonal thanks to each reduction basis normalization. Coupling blocks are not simplified for the projection and are usually full. The final system size is directly the sum of each reduction basis size,

$$\begin{cases} \mathcal{M}_{ii} = \phi_i^T \hat{M}_{ii} \phi_i = \mathcal{I} \\ \mathcal{M}_{ij} = \phi_i^T \hat{M}_{ij} \phi_j \qquad \text{full if } j \in \mathcal{P}_i \end{cases} \tag{2.46}$$

The only solution to reduce coupling computation times is to perform substructure reduction before coupling. An information loss thus occurs, whose control is not as easy as seemed in the first developments of Kaplan [28], and is still an open research field. Gao *et al.* [55, 56] assesses the difficulty of choosing frequency cutting strategies to ensure good convergence of the solution of a given frequency band of interest. In a different context Kammer *et al.* in [57, 58] develop selection techniques for fixed-interface component modes through the computation of an effective mass matrix, which may be of interest.

2.4.3 Practical implementations and actual issues

From the concepts presented in section 2.4.2, practical implementation strategies must be derived. The actual number of operations to perform becomes very large when the front elimination tree shows more than a few levels. The way of sequencing the operations is not unique and very critical. This leads to several algorithm philosophies classically *Lazy* ones or *Greedy* ones [28]

The global procedure can be split in 5 phases, as described in algorithm 1. Step 2, the recursive projection basis $\tilde{T}$ is the base of all AMLS variants.

Phase 3 is critical as the mass projection cost is very sensitive to the implementation. Blockwise operations are exploited but the *choregraphy* of the block computations is critical. The reduction bases are eventually computed in this phase, with several possible methods. The resolution in phase 4 is performed using a standard Lanczos algorithm [41, 51] briefly presented in Appendix B. The solution is then restituted.

Input: M, K, DOF
1. Graph partitioning and matrix reordering
2. Multi-level static condensation
3. Matrices projection
4. Resolution
5. Restitution

Output: $\mathcal{K}, \mathcal{M}, \phi_j, \omega_j, T$

Algorithm 1: AMLS procedure

The algorithm basically performs an iterative static condensation of the children blocks on their parent block. A specific tree path must then be followed such that all children of a given vertex must have been condensed before processing it. Unlike most of the literature, it is here chosen to present algorithms not dissociating leaf/separator cases. The conventions detailed in the following allow compact algorithms close to actual coding.

The following algorithms are written for the treatment of vertex i from the elimination tree. The vertex sequence is a *tree traverse* ensuring that all children of vertex i have been treated before. The children of i are described by the set $\mathcal{C}_i$ and contains all vertices accessible from i to the leaf levels. The parents of i are described by the set $\mathcal{P}_i$ and contains all successive parents of i up to the root level. If $\mathcal{P}_i = \emptyset \Leftrightarrow$ *'i is the root'* or $\mathcal{C}_i = \emptyset \Leftrightarrow$ *'i is a leaf'* the corresponding operations are not performed. The tree branch of node i, containing all its parents and children is noted $\mathcal{B}_i = \mathcal{P}_i \cup \mathcal{C}_i \cup \{i\}$

The complete reduction basis written in (2.44) requires the computation of each substructure modes, *i.e.* all $\hat{M}_{ii}$ and $\hat{K}_{ii}$ blocks. The static condensation operations needed makes this equivalent to computing all $\left[\hat{M}\right]$ and $\left[\hat{K}\right]$. A direct two step procedure would then decompose the matrix projection phase 3 from algorithm 1 as in equation (2.43) with a full projection of $[M]$ and $[K]$ before reducing the system. Such procedure is practically irrelevant as $\left[\hat{M}\right]$ contains potentially large full blocks and is generally much less sparse than the original matrices. Large industrial models at the core of the solver application target would then be impossible to handle.

A reasonable way of tackling the model reduction is to perform substructure reduction on the fly and never generate $\left[\hat{M}\right]$ by directly computing $[\mathcal{M}]$ blocks. The computation organization can then follow several paths with varying performances, with classical implementation in [28]. The computation organization is presented as a set of operations that will be performed for each vertex of the elimination tree.

A natural way of exploiting equations (2.33)-(2.36) is a class of so-called *lazy algorithms* where a minimum of operations are proceeded at each substructure. These algorithms are said to be *left-looking* as children/parent coupling terms are finalized at the parent level, thus looking backwards regarding the tree path followed.

The idea is to apply almost directly formulae (2.35) and (2.36) in an iterative way without considering potential simplifications, yielding the operation sequence when treating a vertex i,

$$
\begin{array}{lll}
1. & (\phi_i, \omega_i) \,/ \hat{K}_{ii}\phi_i = \omega_i^2 \hat{M}_{ii}\phi_i & \\
2. & \tilde{T}_{ij} = -\hat{K}_{ii}^{-1}\hat{K}_{ij} & \forall j \in \mathcal{P}_i \\
3. & \hat{K}_{jj} \leftarrow \hat{K}_{jj} - \tilde{T}_{ij}^T K_{ii} \tilde{T}_{ij} & \forall j \in \mathcal{P}_i \\
4. & \hat{K}_{jk} \leftarrow \hat{K}_{jk} + \tilde{T}_{ij}^T \hat{K}_{ik} & \forall (j,k) \in \mathcal{P}_i \times \mathcal{P}_j \\
5. & \hat{M}_{ij} \leftarrow \hat{M}_{ij} + \sum_{k \in \{i\} \cup \mathcal{P}_i} \hat{M}_{ik} \tilde{T}_{kj} & \forall (j,k) \in \mathcal{P}_i \\
6. & \hat{M}_{kj} \leftarrow \hat{M}_{kj} - \tilde{T}_{ik}^T \left(\sum_{l \in \mathcal{C}_i} M_{li}^T \tilde{T}_{lj} + \sum_{l \in \{i\} \cup \mathcal{P}_i} \hat{M}_{il} \tilde{T}_{lj} \right) + \hat{M}_{ik}^T \tilde{T}_{ij} & \forall (j,k) \in \mathcal{P}_i \times \mathcal{P}_i \setminus \mathcal{P}_j \\
7. & \hat{M}_{ii} \leftarrow \tilde{T}_{ii}^T \hat{M}_{ii} \tilde{T}_{ii} & \\
8. & \hat{K}_{ii} \leftarrow \tilde{T}_{ii}^T \hat{K}_{ii} \tilde{T}_{ii} & \\
9. & \hat{M}_{ki} \leftarrow \tilde{T}_{kk}^T \hat{M}_{ki} \tilde{T}_{ii} & \forall k \in \mathcal{C}_i \\
10. & \tilde{T}_{ki} \leftarrow \tilde{T}_{ki} \tilde{T}_{ii} & \forall k \in \mathcal{C}_i
\end{array}
\tag{2.47}
$$

Phases 2 and 3 of algorithm 1 are coupled such that the recursive projection basis is computed during the projection operations.

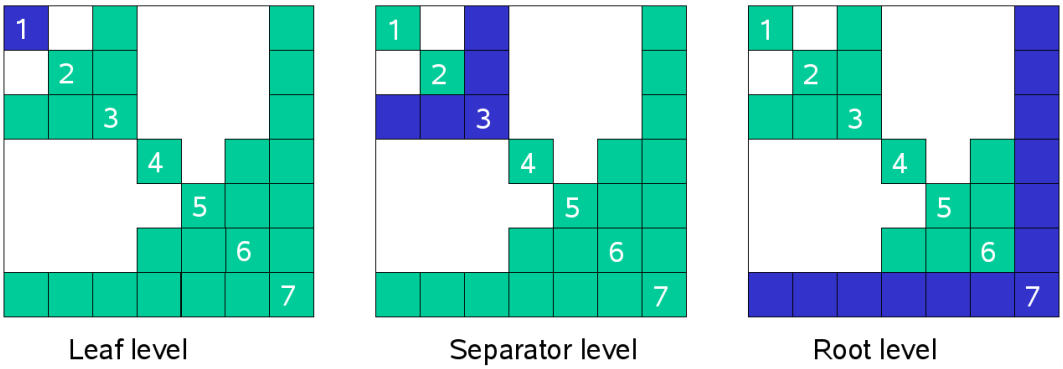


Figure 2.42: Illustration of the lazy algorithm philosophy. Example based on the plate example of figure 2.38. In blue, completed blocks regarding the contribution of the vertex treated (from left to right: a leaf, separator: and root vertex)

Figure 2.42 shows the operation sequence performed from equations (2.47). A minimal number of operations are performed at each operation, so that reduction is only performed for leaf levels, and coupling terms are performed at the separator level, the left-looking characteristic is then highlighted.

Original lazy algorithms were tested as it provides a better product order for left/right mass projections ($T^T MT$ terms). The number of needed operations is however preponderant on the order they are proceeded which penalizes such algorithms. In particular, computing mass coupling terms from the children in equation (2.47)-(5-6) is a very intensive phase.

A second class of *greedy algorithms* performs all possible operations at a given step, these are also said to be *right looking* as children/parents coupling terms are finalized at the child level. The illustration of operation sequence is given in figure 2.43.

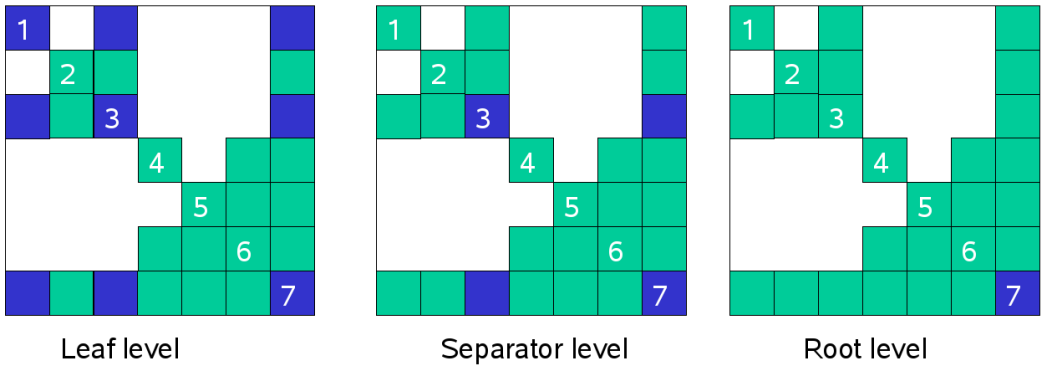


Figure 2.43: Illustration of the greedy algorithm philosophy. Example based on the plate example of figure 2.38. In blue, completed blocks regarding the contribution of the vertex treated (from left to right: a leaf, separator, and root vertex)

There are substantial advantages by doing so as left projections can be proceeded earlier thus minimizing the working memory and reducing matrix sizes for some products. Simplifications in equation (2.36) can eventually be exploited by noting that

$$\tilde{T} = \begin{bmatrix} \mathcal{I} & \cdots & -D_{11}^{-1}\hat{K}_{1i} & \cdots & -D_{11}^{-1}\hat{K}_{1N} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & \mathcal{I} & \cdots & -D_{ii}^{-1}\hat{K}_{iN} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & \mathcal{I} \end{bmatrix} = \prod \tilde{T}^i \tag{2.48}$$

where $\tilde{T}^i$ represents the projection matrix of substructure i on the full model, and therefore only contains coupling terms between i and blocks in $\mathcal{P}_i$.

$$\tilde{T}^i = \begin{bmatrix} \mathcal{I} & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & \mathcal{I} & \cdots & -D_{ii}^{-1}\hat{K}_{iN} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & \mathcal{I} \end{bmatrix} \tag{2.49}$$

Projection on $\tilde{T}^i$ can be written following the notations of (2.42) and (2.44) for any higher level coupling term, yielding

$$\hat{M}_{jk\ \in\mathcal{P}_i\times\mathcal{P}_i} = M_{jk} + M_{ij}^T\tilde{T}_{ik} + \tilde{T}_{ij}^T M_{ik} + \tilde{T}_{ij}^T M_{ii}\tilde{T}_{ik} \tag{2.50}$$

Equation (2.36) can then be replaced by

$$\hat{M}_{ij} = M_{ij} + \sum_{k\in\mathcal{C}_i}\left(M_{ki}^T\tilde{T}_{kj} + \tilde{T}_{ki}^T M_{kj} + \tilde{T}_{ki}^T M_{kk}\tilde{T}_{kj}\right) + \sum_{k\in\mathcal{C}_j}\hat{M}_{ik}\tilde{T}_{kj} \tag{2.51}$$

The sequence (2.47) is thus modified only for the mass coupling terms, in sequence (2.52). Two advantages are sought, seen in phases (2.52)-(5-7). Besides needing less computation terms, projection of the mass coupling terms are treated early enough in (2.52)-(6) to reduce matrix block sizes before all computations are performed.

$$\begin{array}{lll}
1. & (\phi_i,\omega_i)\,/\hat{K}_{ii}\phi_i = \omega_i^2\hat{M}_{ii}\phi_i & \\
2. & \tilde{T}_{ij} = -\hat{K}_{ii}^{-1}\hat{K}_{ij} & \forall j\in\mathcal{P}_i \\
3. & \hat{K}_{jj} \leftarrow \hat{K}_{jj} - \tilde{T}_{ij}^T K_{ii}\tilde{T}_{ij} & \forall j\in\mathcal{P}_i \\
4. & \hat{K}_{jk} \leftarrow \hat{K}_{jk} + \tilde{T}_{ij}^T\hat{K}_{ik} & \forall(j,k)\in\mathcal{P}_i\times\mathcal{P}_j \\
5. & \hat{M}_{jk} \leftarrow \hat{M}_{jk} + M_{ij}^T\tilde{T}_{ik} + \tilde{T}_{ij}^T M_{ik} + \tilde{T}_{ij}^T M_{ii}\tilde{T}_{ik} & \forall(j,k)\in\mathcal{P}_i\times\mathcal{P}_j \\
6. & \hat{M}_{ij} \leftarrow \phi_i^T\left(\hat{M}_{ij} + \hat{M}_{ii}\tilde{T}_{ij}\right) & \forall j\in\mathcal{P}_i \\
7. & \hat{M}_{jk} \leftarrow \hat{M}_{jk} + \hat{M}_{ji}\tilde{T}_{ik} & \forall(j,k)\in\mathcal{C}_i\times\mathcal{P}_i \\
8. & \hat{M}_{ji} \leftarrow \hat{M}_{ji}\tilde{T}_{ii} & \forall j\in\mathcal{C}_i \\
9. & \hat{K}_{ii} \leftarrow \tilde{T}_{ii}^T\hat{K}_{ii}\tilde{T}_{ii} & \\
10. & \hat{M}_{ki} \leftarrow \tilde{T}_{kk}^T\hat{M}_{ki}\tilde{T}_{ii} &
\end{array} \tag{2.52}$$

Such algorithm is more efficient, and was retained by Kaplan in [28]. Since all terms are propagated from the child level to the parents level as early as possible, no backward modification is necessary, excepted modal projections. Although these algorithms are said in the literature to provide more performance for out-of-core computations, the observations made with SDT for models of reasonable sizes as the real brake model provided by Bosch, are not as convincing.

The costliest operation is the projected mass computation, which requires 3 times the cost of the projected stiffness [28]. Such remarks date from 2001 and must be put into perspective with the evolution of computers technology. Great performance differences (over 100%) were detected by Kaplan between 2 machines, either a workstation or a cluster.

From the solver point of view, where actual requirements have shifted from FRF computation of super-large models to the optimization of real eigenvalues resolution for large models, performance variations are a problem. The use of parallelization should nevertheless yield computation performance results below classical Lanczos performance for systems over a few hundred thousand DOF.

The literature available mainly focuses on the algorithms presented in this section, thus taking for granted the iterative representation of the condensation basis. The reality of LDL^T decomposition is however richer than the iterative decomposition presented in section 2.4.2. Indeed, the full condensation basis $[T]$ can be obtained directly by noting that

$$[T] = \left[\tilde{T}\right]^{-1} \tag{2.53}$$

From equation (2.53), a wider range of applied algorithms can be imagined. Direct block projections at any level can indeed be performed, which will relax the mass projection cost constraints. It can also be noticed that the performance of the LDL^T decomposition algorithms used at leaf level is also critical and that many libraries exist for that purpose.

2.5 Subspace regularization methods for contacting interfaces

The present chapter focuses on assemblies by contact on surfaces of complex mechanical parts. The connections at interfaces were assumed to be related to interaction matrices which will be analyzed here, while more details on the implementation of contact will be presented in section 3.2.4.

External modeling and model generation constraints usually found in industry (automatic mesh generation at the component level, mesh provided by contractors, *etc*) often lead to non-conforming meshes for the contacting interfaces. The definition of master/slave surfaces is also a question that is often with no clear answer, in particular with partially covering surfaces due to large displacements or the presence of holes.

The problem of implementing bilateral coupling, or its generalization for contact, for non-conforming meshes has been widely studied in the finite element [59, 60, 61] and component mode synthesis [62, 40] literature. The strategy presented here is based on the notions of pseudo compatibility introduced by Ben Dhia and Balmes [48]. It is here extended to tridimensional meshed volumes. The need to use specific integration rules for interfaces, stated by Ben Dhia [34] is illustrated.

Section 2.5.1 illustrates the problem and proposes a measure of non conforming interface compatibility which can be taken as a quality indicator. The measure is directly exploited in section 2.5.2 to regularize contact solutions by describing interface fields as a linear combination of quasi-compatible shapes, through either primal or dual formulations. Section 2.5.3 eventually provides applications to an academic model showing large mesh refinement differences and to an industrial pad/disc model, showing large non compatible interfaces.

2.5.1 A measure of interface compatibility

A compatibility measure is an indicator of a discretized interface ability to transmit a field from one surface to the other. For a general non conforming interface, not all field distributions can be equally represented by both surfaces. A measure of compatibility thus needs to evaluate differences between the representations of a given field on both surfaces.

The usual finite element formulation of penalized bilateral contact can yield poor results in the case of non conforming interface meshes, especially when mesh sizes differ in contacting surfaces and partial element coverage is present - *e.g.* due to holes or global motion. The academic illustration presented in figure 2.44 is a contact between two cubes in vertical translation. The lower cube is clamped at his base, drilled and finely meshed. The upper cube is plain, coarsely meshed and a pressure is applied to its top surface.

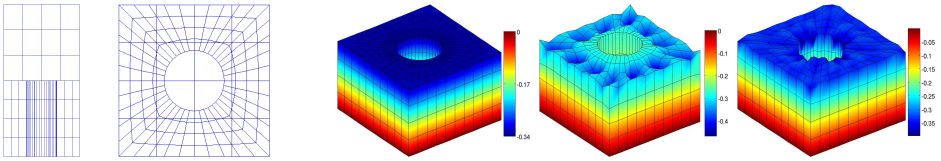


Figure 2.44: Example of incompatible mesh with partial covering. Reference solution is plotted in the middle, against the 1 (right) and 4 (far right) point rule integration solution using the upper cube as master

Figure 2.44 presents displacement results obtained for different formulations. Using the coarse upper cube as the master means that a lot of slave nodes are not matched (or not *seen* by any master contact point). The richer shapes of the refined cube cannot be represented on the coarse cube and are therefore not seen nor constrained by the contact formulation on the coarse cube. This pattern is close to hourglass modes observed for under-integrated elements.

Introducing very rich integration avoids oscillations but can significantly increase computational costs and leads to excessive stiffness, or locking, when zero gap is enforced strictly [48]. The following sections will introduce an energy function for the gap and show how this can be used to quantify and possibly relax compatibility of displacements on the interfaces.

Contact between two solids Ω_1 and Ω_2, is defined between two surfaces Γ_1 and Γ_2. The compatibility is computed on the subset of effective contact Γ of $\Gamma_1 \times \Gamma_2$, as shown figure 2.45.

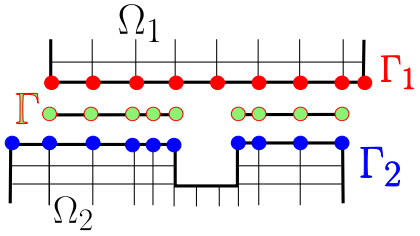


Figure 2.45: Definition of a domain Γ for the compatibility computation

Any contact formulation needs to evaluate the gap field between two surfaces (section 3.2.4), which is defined for a displacement $u(q_1)$ of Γ_1 and $u(q_2)$ of Γ_2, where q_1 (resp. q_2) is the discretized displacement on Γ_1 (resp. Γ_2) as

$$\{g(u(q_1), u(q_2))\} = \{u(q_1) - u(q_2)\}^T \{n\} \tag{2.54}$$

The first step is to define a scalar product on $\Gamma_1 \times \Gamma_2$ that is a norm for the gap. The strain energy of a penalized contact with a uniform contact stiffness density k is chosen here. For a displacement u, this energy is given by

$$E_p = \int_\Gamma kg\left(u(q_1), u(q_2)\right)^2 dS \tag{2.55}$$

This scalar product must be approximated as precisely as possible to avoid poor conditioning issues, as stated by Ben Dhia [34]. The computation strategy chosen here is an automated Delaunay triangulation over the subset Γ of $\Gamma_1 \times \Gamma_2$ in effective contact. The scalar product is computed by numerical quadrature from the mesh of Γ

$$E_p = \sum_{x_i} kw_i J(x_i) g_i \left(u_i(q_1), u_i(q_2)\right)^2 \tag{2.56}$$

In practice (section 3.2.4), the gap is computed on N_Γ integration points (Gauss) and linearly linked to the normal displacement observation matrices $\left[C_N^1\right]_{N_\Gamma \times N_{\Gamma_1}}$ (resp. $\left[C_N^2\right]_{N_\Gamma \times N_{\Gamma_2}}$) of Γ_1

(resp. Γ_2) on Γ. This yields the discretized gap formulation

$$\{g\}_i = \left[C^1_{N_i}\right]\{q_1\} - \left[C^2_{N_i}\right]\{q_2\} \tag{2.57}$$

The contact strain energy is then of the form

$$E_p = \left\{ \begin{array}{c} q_1 \\ q_2 \end{array} \right\}^T \left[\begin{array}{cc} C^1_{N_i} & -C^2_{N_i} \end{array} \right]^T \left[\backslash k w_i J(x_i) \backslash\right] \left[\begin{array}{cc} C^1_{N_i} & -C^2_{N_i} \end{array} \right] \left\{ \begin{array}{c} q_1 \\ q_2 \end{array} \right\} \tag{2.58}$$

which defines the scalar product on Γ. The scalar product matrix is noted $[A]$ and is written by blocks

$$E_p = \{q\}^T [A] \{q\} = \left\{ \begin{array}{c} q_1 \\ q_2 \end{array} \right\}^T \left[\begin{array}{cc} A_{11} & A^T_{21} \\ A_{21} & A_{22} \end{array} \right] \left\{ \begin{array}{c} q_1 \\ q_2 \end{array} \right\} \tag{2.59}$$

In [48] the so-called $(1-\epsilon)$-compatibility is computed by considering the norm difference between a displacement q_1 of an interface Γ_1 and its projection to the facing interface $\pi^1_2\{q_1\} \in \Gamma_2$. The projection, defined as the vector of Γ_2 that minimizes the gap energy, verifies

$$[A_{21}]\{q_1\} - [A_{22}]\pi^1_2\{q_1\} = 0 \tag{2.60}$$

This suggests a formulation of the $(1-\epsilon)$-compatibility as a Rayleigh quotient,

$$C^1_2(\{q_1\})^2 = \frac{\|\pi^1_2\{q_1\}\|_{A_{22}}}{\|\{q_1\}\|_{A_{11}}} = \frac{\{q_1\}^T [A_{21}]^T [A_{22}]^{-1} [A_{21}] \{q_1\}}{\{q_1\}^T [A_{11}] \{q_1\}} \tag{2.61}$$

Such formulation raises robustness issues as partial element covering tends to yield ill conditioning of $[A_{11}]$ and $[A_{22}]$. The formulation presented here computes the singular value decomposition (SVD) of $[A]$ (rather than $[A_{21}]$ suggested in [48]). As $[A]$ is symmetric,

$$[A] = \left[\ldots \left\{ \begin{array}{c} u^1_i \\ u^2_i \end{array} \right\} \ldots \right] \left[\backslash \sigma_i \backslash\right] \{U\}^T \tag{2.62}$$

The SVD generates a displacement basis $\{u_i\}$ of the coupled interface with the level of gap energy generated σ_i. Low singular values thus characterize compatible interface displacements. The compatible displacement fields obtained are then known to be represented on both surfaces well enough - which will be quantified by a tolerance ϵ. A new definition of interface compatibility is suggested as

$$C^1_2(\epsilon) = \frac{card\left\{i / \frac{\sigma_i}{\max_i \sigma_i} \leq \epsilon\right\}}{\min_{i=1,2}(card(\Gamma_i))} \tag{2.63}$$

It is useful to note that the gap observation (2.57) can be performed with the full contact set. Unmatched contact DOF will generate zero terms on the diagonal of $[A]$. No conditioning issue affects the SVD computation which would output in that case a null (to numerical precision) singular value associated with shapes showing unmatched points displacements only. The description of the unmatched point displacements is then a full recombination of the finite element basis. The refinement of Γ is nevertheless necessary for a good description of the non spurious contact points.

2.5.2 Using (1-ϵ) vector pairs to solve contact problems

The compatibility measure through an SVD generates couples of quasi-compatible displacements on both sides of an interface. These represent bases that can be exploited by subspace generation methods in which the non-compatible shapes can be controlled or removed, thus regularizing the contact solution.

The first idea is a primal formulation which will constrain the interface displacement to the subspace of quasi-compatible displacements, namely T_ϵ^1 and T_ϵ^2. These bases contain the shapes under the tolerance ϵ, using definition (2.63). The model DOF are sorted to segregate the interface DOF (q_{i1} and q_{i2}) from the rest (q_c), thus allowing the projection to the generalized quasi-compatible interface DOF ($q_{\epsilon 1}$ and $q_{\epsilon 2}$). Regular contact force distributions are then a consequence of the imposed displacement regularity.

$$\left\{ \begin{array}{c} q_c \\ q_{i1} \\ q_{i2} \end{array} \right\} = \left[\begin{array}{ccc} \mathcal{I} & 0 & 0 \\ 0 & T_\epsilon^1 & 0 \\ 0 & 0 & T_\epsilon^2 \end{array} \right] \left\{ \begin{array}{c} q_c \\ q_{\epsilon 1} \\ q_{\epsilon 2} \end{array} \right\} = [T_\epsilon] \left\{ \begin{array}{c} q_c \\ q_{\epsilon 1} \\ q_{\epsilon 2} \end{array} \right\} \tag{2.64}$$

The gap observation is consequently projected and can be written

$$\{g\} = \left[C_{Ni}^1 \left[T_\epsilon^1 \right] \right] \{q_{\epsilon 1}\} - \left[C_{Ni}^2 \left[T_\epsilon^2 \right] \right] \{q_{\epsilon 2}\} \tag{2.65}$$

Exact compatibility (zero gap on Γ) is known to generate locking issues, as discussed for example in [62, 48]. Introducing quasi-compatible displacements relaxes the constraint while avoiding large relative displacements.

In a dual formulation, the projection on the (1-ϵ)-compatible displacement pairs of the contact forces is assumed to be zero. This clearly avoids local stress concentrations associated with locking but also leads to larger gaps. Dual formulations also have the advantage of preserving the use of physical displacement DOF.

In general penalized contact approaches, the load is a non linear function of the gap. Exponential functions are in particular used for brake squeal studies, section 3.2.4. In a dual formulation, locking is avoided but the gap is no longer regular. It is thus proposed to use a regularized gap observation on Γ to compute contact forces.

For a surface Γ_n the regularized displacement is a linear combination of the trace on Γ_n of the compatible vectors, $[T_\epsilon^n]$

$$\{q_n^\epsilon\}_{N_{\Gamma_n} \times N_\epsilon} = [T_\epsilon^n]_{N_{\Gamma_n} \times N_\epsilon} \{\alpha_n\}_{N_\epsilon} \tag{2.66}$$

where $\{\alpha_n\}$ is a solution of the minimization problem

$$\{\alpha_n\} = \min_{\{\alpha\}} \| [T_\epsilon^n] \{\alpha\} - \{q_n\} \|_{\mathcal{K}} \tag{2.67}$$

This defines a pseudo-inverse, as the solution searched is the closest to $\{q_n\}$ possible, relatively to a norm $\mathcal{K}$. The choice of $\mathcal{K}$ is fully open. It cannot be a restriction of $[A]$ due to conditioning problems, but considering a norm in displacement or strain on Γ_n seems meaningful. The identity was chosen here so that the modified gap observation is

$$\{g\}_i = \left[C_{Ni}^1 \left[T_\epsilon^1 \right] \left[T_\epsilon^1 \right]^+ \right] \{q_1\} - \left[C_{Ni}^2 \left[T_\epsilon^2 \right] \left[T_\epsilon^2 \right]^+ \right] \{q_2\} \tag{2.68}$$

Projection of both sides of the interface can be considered. In master/slave configurations, a slave surface only projection is sufficient to suppress the observation of non quasi-compatible displacements.

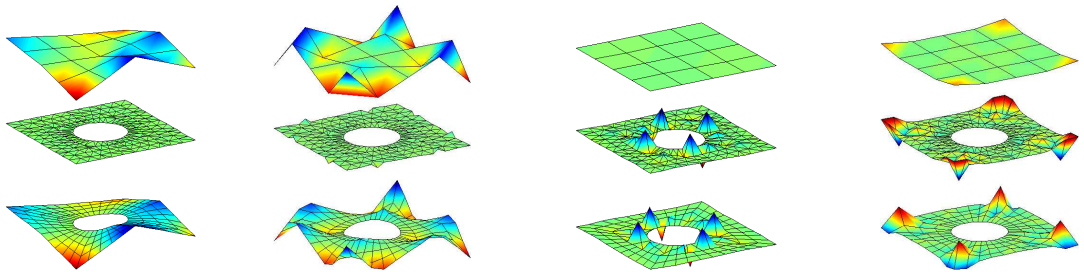
Obtaining proper quasi-compatible modes requires a good representation of the scalar product presented in section 2.5.1. Relatively rich integration rules on both surfaces are therefore needed for these computation steps. However, the selection of quasi-compatible modes typically restricts the number of needed interface degrees of freedom significantly. The number of integration points needed for contact can consequently be decreased as information redundancy occurs. This is particularly interesting in vibration studies for which the contact surface is assumed constant during long time simulations, so that coarser rules will decrease computation time.

The proposed strategy is to use a fine integration rule to compute $[A]$ and $[T_\epsilon^i]$ defined in (2.61) and (2.62), then to switch to a coarser rule for the evaluation of the regularized gap (2.68) and resulting contact loads. Such under-integration could be optimized by the use of the Orthogonal Maximum Sequence method developed by Balmes [63] to create an optimal set of contact points associated to the quasi-compatible shapes to be observed.

2.5.3 Illustrations

The concepts of the $(1-\epsilon)$-compatibility are applied to the academic model which features great mesh refinement differences, and to an industrial model featuring a relatively large interface. Regularization patterns are presented on the first case while under-integration strategies in the scope of an improvement of dynamical behavior are discussed on the latter model.

This section illustrates the resolution method presented to the drilled cube example, using the upper cube as contact master, which yields poor results when no regularization is considered. The first step to the resolution is the creation of Γ, and the SVD computation of $[A]$, as plotted in figure 2.46.



(a) Shape #1, $\sigma = 4.2.10^{-18}$ (b) Shape #22, $\sigma = 5.3.10^{-4}$ (c) Shape #96, $\sigma = 7.9.10^{-2}$ (d) Shape #166, $\sigma = 1.2.10^{-1}$

Figure 2.46: Shapes resulting from the SVD of $[A]$ plotted on Γ_1 and Γ_2. The gap is computed and plotted on the Gauss points of Γ

The gap is computed on the Gauss integration points of Γ for which a Delaunay triangulation is performed for visualization purposes. The compatible shapes shown in figure 2.46a or 2.46b correspond to very small gaps. The incompatible shapes in figure 2.46c and 2.46d show a high singular value (8% and 12%) and correspond to a large gap visible on Γ.

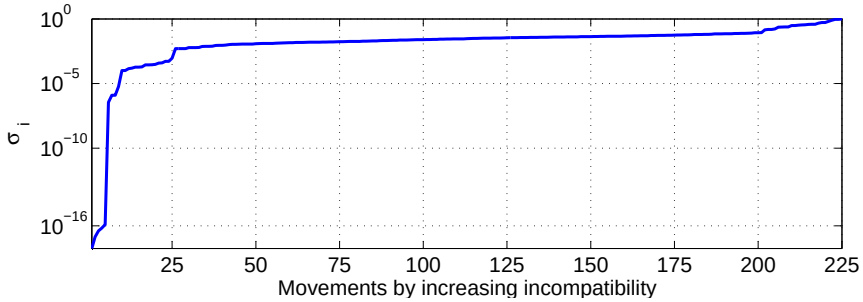
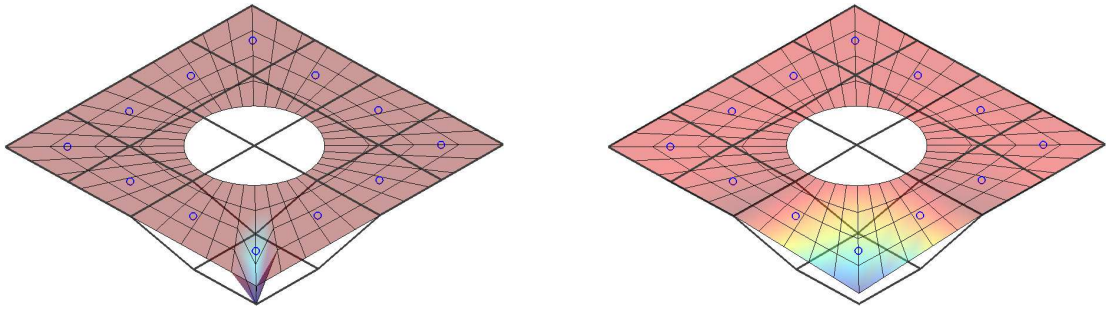


Figure 2.47: Singular values of $[A]$ in increasing incompatibility

In this problem, 25 quasi-compatible shapes at most exist since this is the number of DOF of the coarse interface. A clear jump can be seen in figure 2.47 after this threshold.

To illustrate the regularized gap approach, the projected observation $\left[C_N^1\right]\left[T_\epsilon^1\right]\left[T_\epsilon^2\right]^+\Big]$ is plotted and compared to $\left[C_N^1\right]$ in figure 2.48. The observation is the dual of the contact pressure as they are directly related by the pressure-gap contact law. The regularization method

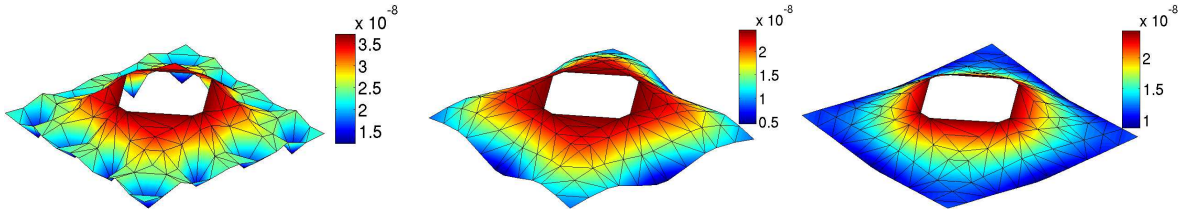
clearly distributes the gap observation to the contact point vicinity, instead of having a point-to-matched-point observation.



(a) Initial observation (b) Regularized observation

Figure 2.48: Initial and regularized observation for a given Gauss integration point (o) of the master surface

Figure 2.49 plots gap results obtained for the center (one) point integration rule. The basic results are very poor as said earlier, and both primal and dual regularized solutions are satisfying.



(a) Center rule (b) Center Dual Reg. (c) Center Primal Reg.

Figure 2.49: Gap observations obtained using the center integration rule for contact, observed on the rich 3x3 rule Gauss points Delaunay triangulation, $\epsilon = 10^{-5}$

Basic solutions provide poor contact fields, using a richer integration rule for that case is however an improvement. Indeed the approached contact area seen by the covered contact points increases by 25% when switching from the center point rule to the 2x2 points rule.

The primal resolution directly constraints displacements but can raise locking issues. No such pattern is observed for the cube and direct displacement restriction yields in this case better contact pressure fields. It is however costlier to implement as the interface projection would typically require DOF reordering prior to solve.

State-of-the-art brake models used for the industrial application are eventually tested. The system is meshed automatically by component using 10 node tetrahedron elements, resulting in non conforming interfaces. Figure 2.50 shows the components retained, the disc and a pad with its backplate and lining. Pressure is applied using the trace of the caliper fingers on the backplate.

Static responses are properly handled here using a rich contact integration rule (even over-integration). For applications in dynamics, one seeks to compute small oscillations around a bilateral contact state. The measure of incompatibility introduced in the paper shows that although the mesh sizes are similar, compatibility is not very good. The combination of incompatible mesh and rich integration, then leads to *locking* issues which will be illustrated at the end of this section.

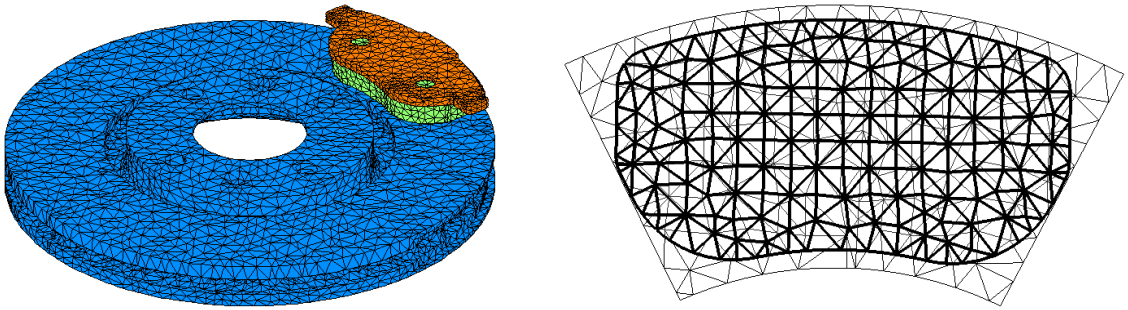


Figure 2.50: Pad/Disc model and superimposed interface discretization

Applying the concepts presented leads to the compatible singular values and associated compatibility measure shown in figure 2.51. The first 10% of compatible shapes is free movements for the uncovered nodes of the partially covered elements. A threshold ϵ value is seen before more quasi-compatible shapes are found. More than 70% of shapes are above a $\epsilon = 10^{-4}$ threshold. Thus although mesh sizes are similar, the compatibility is low. Enforcing strict displacement equality is thus expected to induce locking.

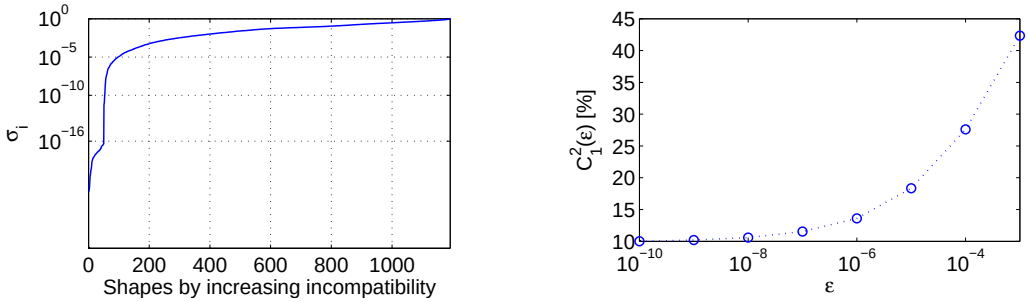
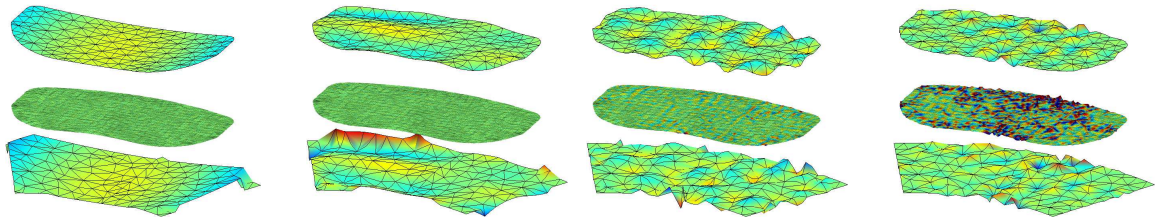


Figure 2.51: Left: Singular values of $[A]$ - Right: (1-ϵ)-compatibility as function of ϵ

Sample shapes resulting from the SVD are plotted in figure 2.52. Shapes are plotted on Γ_1 and Γ_2 which are the upper and lower surfaces. The gap between both surfaces displacement is plotted on Γ shown at the intermediate height.



(a) Shape #52, $\sigma = 1.1.10^{-11}$

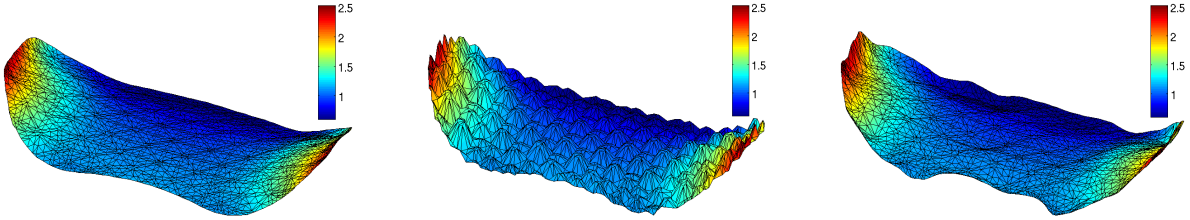
(b) Shape #73, $\sigma = 1.1.10^{-6}$

(c) Shape #150, $\sigma = 1.2.10^{-4}$

(d) Shape #217, $\sigma = 8.2.10^{-4}$

Figure 2.52: Sample quasi-compatible shapes. Γ is represented as a Delaunay triangulation of its Gauss points

The resolution of a contact problem with an exponential law is here satisfying using a rich integration rule (6 points for a 6 nodes triangle), but requires 1824 contact points. A center point integration rule uses only 394 points but results in strong contact irregularities, as illustrated in figure 2.53.



(a) 6 point rule (b) Center point rule (c) (1-10^{-5}) rule

Figure 2.53: Basic contact pressure solution [MPa] for rich rule (6 point), poor rule (1 point), and regularized 1 point observation using shapes computed with the 6 points rule

Figure 2.53c clearly illustrates that the regularized gap function avoids large oscillations and yields smoother results when a small ϵ (10^{-5}) is considered. The pressure does not however exactly converge towards the rich 6 point per element rule. Further work on the reduction of the number of contact points will thus need to address the relation between the localization of contact points and the inherently non-local nature of (1-ϵ) compatible deformations.

As well documented in the literature, over-integration is a good solution to solve static contact problems. In dynamics, enforcing permanent contact exactly for incompatible meshes leads to locking. As an illustration, figure 2.54 shows a transfer function from the vertical displacements of a disc point to a pad point. Modes are computed enforcing perfect bilateral contact (no gap) of the contact interface through either the basic contact normal displacement observation or the modified one (equation (2.68)), using $\epsilon = 10^{-3}$.

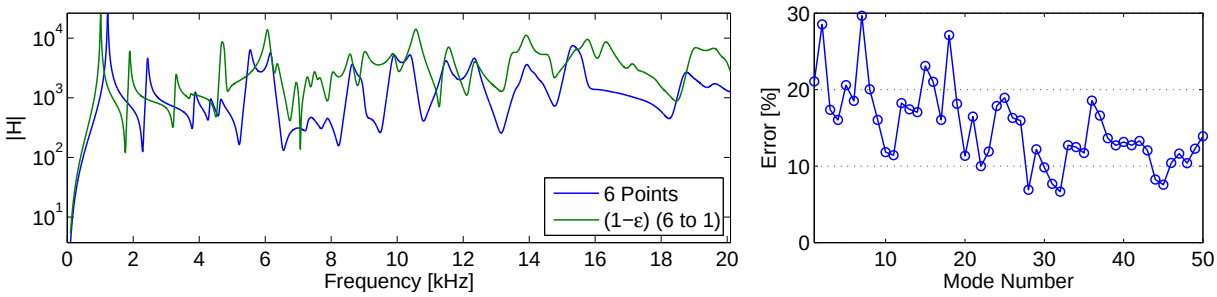


Figure 2.54: Left: Dynamic behavior differences between the use of a 6 point integration rule contact integration law, and the regularization strategy based on a 1 point integration rule for $\epsilon = 10^{-3}$ (42% of quasi-compatible shapes kept). Right: Frequency differences observed at peaks

The over-integrated solution shows positive frequency shifts from 6 to 30%, corresponding to resonance shifts visible in the figure even at low frequencies and to a clear stiffening, or locking, of the structure. The application of interest being brake squeal simulations, the typical frequency range of interest goes up to 16 kHz and the great differences observed in this range must clearly be addressed.

In a penalty based contact, the existence of a tangent stiffness may alleviate some of the problems. Some of the locking may disappear for a soft contact. The non-physical nature of such contact is however not very satisfactory.

The developments of this section are still exploratory. Regarding the industrial stakes and the already complex configuration of a real brake model, the pad disc interface was remeshed for the following thesis applications. The implementation of other non-linear interfaces could however be treated by such technique in perspective works.

3

Contact-friction induced instabilities

Contents

3.1 Introduction

The study of instabilities is of critical interest for industrial applications. More recently instabilities due to friction have been put under the spotlight; these are complex problems involving local vibrations which are responsible for fatigue, loss of control, and noise. In the brake squeal application, the instability is non destructive, the noise generated is however a burden which must be resolved.

The recent improvement in simulation capabilities allows nowadays simulating large assemblies, featuring interacting components. This requires the implementation of contact and friction models in the finite element environment used for mechanical simulations. Section 3.2 thus assesses the implementation of such interaction models, which have produced a lot of literature. The book of Sextro [64] provides a pragmatic description of current applications.

The study of instabilities of mechanical systems is obviously of particular interest due to their potential catastrophic consequences. The study of flutter instability in aeroelastics, in [65] for example amongst so many, is critical for aircraft applications since it can be destructive. For squeal applications, the instabilities are not destructive but the noise generated is very negatively considered. The aircraft brake applications by Sinou *et al.* can become more critical, as presented in [66, 67].

A lot of literature has been produced, briefly presented in section 3.3. Friction in particular has been studied for example by Basseville [68] who studies the stability equilibrium of simple systems. Nguyen [69] gives an insight of stability studies for all types of frictional problems.

3.2 Modeling contact and friction

Contact modeling has been the center of attention for more than a century. Fatigue problems due to large industrial systems (wheel-rail contact for trains, plane-sphere contact for bearings) needed to be assessed for design and failure expertise. Normal contact was studied in particular by Hertz in the last decades of the XIXth century [70] to analytically handle surface contact of elastic bodies. The surfaces were then considered perfectly smooth, which was sufficient for the time being, particularly considering metrology limitations.

The study of friction was rationalized even earlier, already studied by de Vinci in the XVIth century. Its effect was indeed critical to understand due to its resistive character. The dissipations due to friction needed compensation in systems, which required power adjustments. Amontons in 1699 [71] was indeed interested in the origin of resistance in machines and set up the classical formalism of friction laws later confirmed by Coulomb [72], who stated the principles of friction still used nowadays. It was observed the static/dynamic friction transitions, the proportionality between the friction force and normal load in sliding, and the role of roughness expressed as *friction can only come from the gearing of surface asperities*. The static/dynamic transitions where later studied in the 1900's by Stribeck [73] who suggested a more comprehensive approach of the phenomenon. This is still known nowadays as the Stribeck effect, although his studies were targeting different applications and friction conditions.

In the 50's, the work of friction was studied under the Hertz theory by Mindlin [74], and further works were conducted by Johnson [75], and related to wear by Archard [76]. Here experimental consideration and numerical concerns split up, as Signorini in 1933 [77] setup a formalism to treat contact with perfectly plane surfaces, while later in 1966 Greenwood and Williamson [78] proposed the first measures of asperities.

This concept distinction still exists today between pure numericians having formalized ideal surfaces and developing specific resolution algorithms and applied mechanicians who prefer experiment based laws relating documented phenomena, but also requiring parameter identification. In the first case, numerical difficulties arise, due to the loss of solution unicity properties

and non-smooth transitions, which was already noted by Painlevé as early as in 1895 [79], and for example studied by Ballard and Basseville [80] for simple systems.

3.2.1 Ideal Signorini-Coulomb model

The most simple contact representation – at least in formalism – is commonly known as the Signorini-Coulomb law to relate contact and friction. The definition of contact between to solids requires the notion of *gap* giving the distance between the solids considered, and taken as positive if non contact occurs. The notion of *contact normal* naturally follows as giving the way of measuring the gap. As presented in figure 3.1, the gap is measured as the shortest distance between two solids thus along the normal between both solids. It is function of the deformation u of each solid.

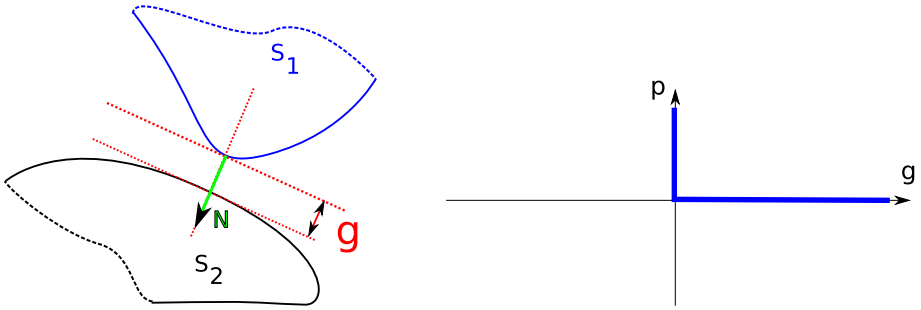


Figure 3.1: Contact normal (N) and gap (g) definition for two solids, and the Signorini contact-pressure law

In case of contact, a reaction force exists between both systems to avoid interpenetration. As contact between three dimensional solids is surfacic, the notion of contact pressure is preferred here, and will be denoted p. It is of course directed by the contact normal. The Signorini conditions thus states for each point in contact

$$\begin{cases} g \geq 0 \\ p \geq 0 \\ (g).(p) = 0 \end{cases} \tag{3.1}$$

It sets exclusion between the gap values and contact pressure values. Two states are then possible, *opened* contact with a strictly positive gap and *closed* contact with a null gap and a strictly positive contact pressure. The law is plotted in figure 3.1.

The ideal Coulomb friction, as originally presented considers a so-called *static* friction model. The formulation distinguishes *sticking* and *sliding* states, depending on the existence of movement between the bodies in contact. If the solids in closed contact do not slide, a reaction force – the friction force – exists between both. This force is in the tangent plane defined as orthogonal to the contact normal N, and is determined by the equilibrium of the whole system. In the case of sliding, the friction force is opposed to the sliding velocity w, and only proportional to the normal load (or contact pressure) through the introduction of a *friction coefficient* μ.

Coulomb friction can be expressed as

$$\begin{cases} \| \{f_T\} \| \leq \mu \| \{f_N\} \| \\ \| \{f_T\} \| < \mu \| \{f_N\} \| \Leftrightarrow \{w\} = \{0\} \\ \| \{f_T\} \| = \mu \| \{f_N\} \| \Leftrightarrow \{f_T\} = \mu \| \{f_N\} \| \frac{\{w\}}{\|\{w\}\|} \end{cases} \tag{3.2}$$

where f_N is the contact force, f_T the friction force, μ the friction coefficient, and w the sliding velocity. The law is plotted in figure 3.2 which represents the friction force as function of the

sliding velocity. For non-sliding states, the friction force is not defined by the law, as only an inequality is set.

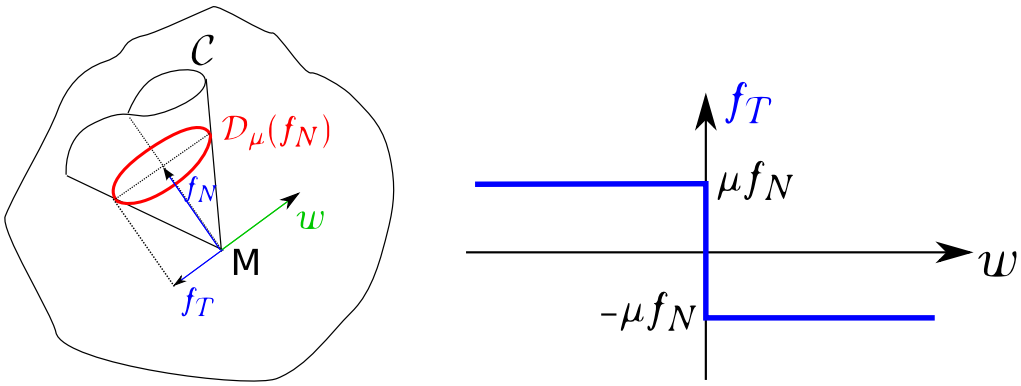


Figure 3.2: Ideal Coulomb friction law and representation of the Coulomb cone

The inequality concept is classically represented by the Coulomb cone shown in figure 3.2. For a given contact force at point M, the reaction force is said inside the coulomb cone $\mathcal{C}_\mu$, and more precisely in the cone section $\mathcal{D}_\mu$, determined by the tangent plane direction and the contact force amplitude.

The Signorini-Coulomb laws given by equations (3.1) and (3.2) raise numerical and analytical issues, since contact and friction forces are not defined for each point of their parameter. Indeed, for zero sliding velocity and zero gap, the forces are unknown and must be determined by other means.

To solve such problem a Lagrangian formulation has to be used, where the contact-friction forces are considered unknown and kept in the system to solve. The system size thus increases as function of the number of contact points, and features inequalities.

The resolution of such systems is not trivial numerically and generated a large range of methods for static and dynamic resolutions. For statics, Moirot [5] proposed a method where an iteration is based on the status assumption of each contact points (open/close gap, sticking/sliding friction) allowing resolving the mechanical equations in a more simple way. De Saxcé and Feng developed the so-called *bipotential* method [81] which proved to be efficient in numerous cases. Several authors developed other mathematical considerations, performing least square resolutions as developed by Haslinger [82], or considering a convex optimization by Acary *et al.* [83]. Alart and Curnier [84] also presented a generalization of the Newton method adapted to Lagrange formulations. This was later improved by Jourdan and Alart through the introduction of Gauss–Siedel like algorithms [85]. The use of augmented Lagrange parameters has been used by Ben Dhia and allowed him to develop an efficient formalism for contact modeling [86].

The aim here is not to detail all these methods but only to present the concepts allowing solving Lagrangian contact-friction, mainly by the transformation of the Signorini-Coulomb equations into projection relations. The finality of such transformation is to allow an iterative resolution of the mechanical contact-friction equations, through an Uzawa algorithm for example.

The Signorini condition can be replaced by

$$\{f_N\} = \mathsf{P}_{\mathcal{R}^-}(\{f_N\} - \rho_N(\{g\} - \{g_0\})) \tag{3.3}$$

where $\mathsf{P}_{\mathcal{R}^-}$ is the projection operator in the real negatives space, and ρ_N is a scalar (or vector) parameter for numerical adjustments.

The Coulomb law equivalent to (3.2) can be reformulated using the notations previously introduced, following the maximum dissipation principle (as presented for example by Jean [87]).

Indeed, there is either a stick state in which case $\{w\} = 0$, or a slip state in which case the friction force is maximum. A test friction force will then verify

$$\forall \{f_T\}^\star \in \mathcal{D}_\mu(f_N), (\{f_T\}^\star - \{f_T\}) \cdot \{w\} \geq 0 \tag{3.4}$$

The formulation (3.4) allows verifying that formulation (3.5) is equivalent to (3.2)

$$\{f_T\} = \mathsf{P}_{\mathcal{D}_\mu(f_N)} \left(\{f_T\} - \rho_T \{w\}\right) \tag{3.5}$$

where $\mathsf{P}_{\mathcal{D}_\mu(f_N)}$ is the projection on the Coulomb cone section operator. The proof of the equivalence of (3.5) and (3.4) does not require a more specific value for ρ_T, which is a numerical adjustment parameter (scalar or vector).

Formulation (3.3) and (3.5) are used in practice as they naturally provide an iterative resolution algorithm. Indeed, given an initial value for f_N and f_T, a correction can be computed through the projection. The correction is so to speak a repositioning of the contact-friction forces in their boundaries.

Using such technique, the mechanical result does not depend on the parameters ρ_N and ρ_T. They must be optimized for the scheme to converge. Indeed setting them too low would slow down the convergence by applying always small corrections to the forces; setting them too high however will cause divergence. It can be noticed that ρ_N is homogeneous to a stiffness and ρ_T to a damping value. Their effect on convergence can then be assessed like for usual modified Newton techniques.

A simple Uzawa algorithm can be setup, as presented in figure 3.3, where a prediction-correction algorithm on the contact-friction forces is used for the mechanical equilibrium resolution, giving new displacement results until convergence.

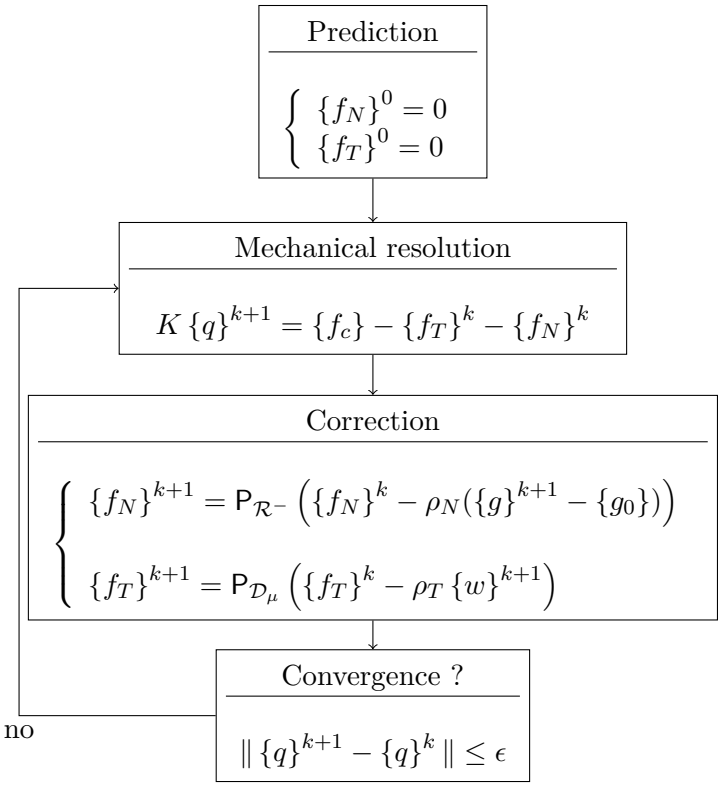


Figure 3.3: Implementation of an Uzawa resolution algorithm for the Signorini-Coulomb contact-friction laws

In practice, the convergence of such algorithm is very sensitive to the parameters ρ_N and ρ_T which can be scalars or vectors. Their values must be chosen in accordance with the mechanical properties of the system, through the spectral radius of the stiffness matrix. The bipotential method for example suggests a specific value for each contact point based on the local stiffness [88].

3.2.2 Functional representation of contact pressure

In addition to the numerical difficulties due to the Signorini-Coulomb formalism, the idealization represented may not always be satisfying. The reality of surfaces is indeed roughness describing the irregularities coming from the asperities of both surfaces. This is conceptualized in figure 3.4.

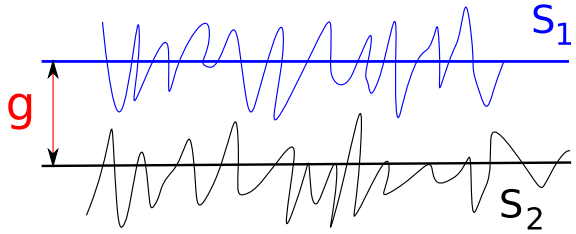


Figure 3.4: Illustraton of roughness against nominal surfaces

Due to the vey small scales of these asperities – around the micrometer – they are obviously not modeled in details. The roughness measures in particular commonly realize statistics on the whole surface, yielding a few parameters describing the surface state. Of course simulation of mechanical components consider nominal surfaces, with no asperities, as represented in figure 3.4.

In the world of asperities averaged by a nominal surface, interpenetration is possible at the macroscopic level, as function of the pressure applied on both surfaces. Indeed, the compression of asperities then increases allowing a greater contact pressure. From macroscopic tests, it is thus possible to obtain a law relating gap to contact pressure.

Possible functional representations of the pressure/gap relation are linear, power series or exponential laws, illustrated in figure 3.5.

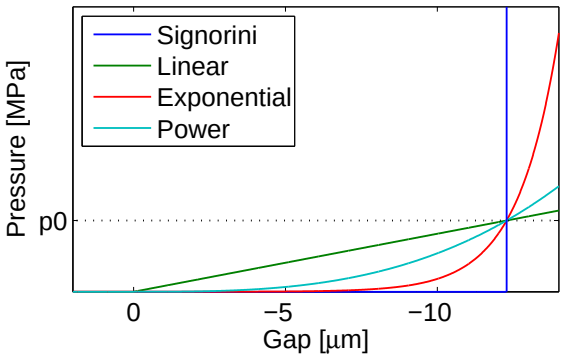


Figure 3.5: Classical contact laws

In practice, the contact pressure is applied to a deformable body. Assuming a certain height for this body, its stiffness k_e can be defined. The measurable displacement is related to pressure by

$$q = g(p) + p\frac{S}{k_e} \tag{3.6}$$

which is illustrated in figure 3.6.

Extreme regimes can easily be seen in equation (3.6). For a very thin body, the stiffness k_e is very high and the deformation will be mostly due to the deformation of asperities. In such case, a pressure/gap relation should be used.

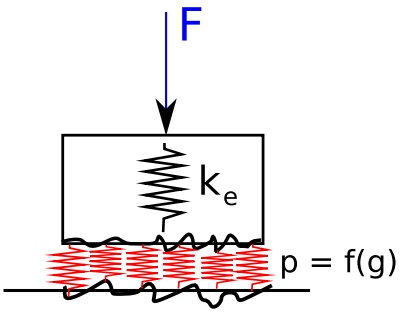


Figure 3.6: Representation of contact between two bodies, body stiffness and asperities

For many more configurations, the gap induced by realistic pressures will be small compared to deformations in the body. This regime is the reason for the widespread use of Signorini conditions. Since the real penetration is small, it cannot be measured accurately against body deformation and assuming it to be zero thus makes more sense. Furthermore, using it would cause major numerical difficulties since the accuracy on gap increments would need to be much higher than the accuracy desired for global motion.

Functional representations of contact pressures are also used in penalization methods. In such case, gaps obtained in the simulation can be seen as numerical errors. A relaxation of the Lagrange constraints is thus performed, authorizing interpenetration whose trade-off is the addition of a force penalizing the violation of the constraint. The best interpretation is obtained energetically, where the relaxed constraint violation is generates a highly energetic term. Consequently, the mechanical solution will only show limited violations of the constraint.

For contact between metals, one can reasonably assume that Signorini conditions should be used. For squeal applications, the pads are significantly softer than the disc so that the more *'physical'* strategy is unclear. One should also note that in finite element simulations, the appropriate model will depend on element size with, functional representations being more appropriate for small elements.

It is noted here that the zero gap value that would be used in a Signorini model does not normally correspond to zero pressure, the contact of the mean smooth surfaces considered in this case corresponds to a state where some asperities are already into contact.

The pressure-gap relationships are defined for each contact point, following definitions presented in table 3.1,

Law	Name	Parameters
$p(g) = -k_c min(0, g)$	Bilinear	k_c
$p(g) = \sum_l p_l g^l$	Power	l, p_l
$p(g) = p_0 e^{-\lambda g}$	Exponential	p_0, λ

Table 3.1: Sample contact laws regularizations through pressure-gap relationships

Using such laws, the contact forces can be condensed on the displacement, making the mechanical resolution much simpler. The contact law parameters must however be identified experimentally which may be difficult, whereas Signorini requires numerical adjustments. Contact mechanism is well assessed, and both Lagrangian and penalization formulations have their adepts.

Questions of parameter sensitivity and numerical simulation robustness will play in favor of regularized contact laws. In particular, Bosch uses experimental exponential contact laws, with parameters p_0 and λ.

3.2.3 Regularized friction models

Friction behavior is more complex, and the Coulomb law, as simple as it is, does not satisfy the needs of everybody.

On the first hand, regularization methods can be employed to relax the Coulomb law for zero sliding velocities. The idea is then to penalize low sliding velocities through the introduction of a slope at the origin, introducing a parameter k_t, so that

$$\begin{cases} \| \{f_T\} \| = k_t \| \{w\} \| & \text{if } \| \{w\} \| < \frac{\mu}{k_t} \| \{f_N\} \| \\ \| \{f_T\} \| = \mu \| \{f_N\} \| & \text{else} \end{cases} \tag{3.7}$$

as shown in figure 3.7.

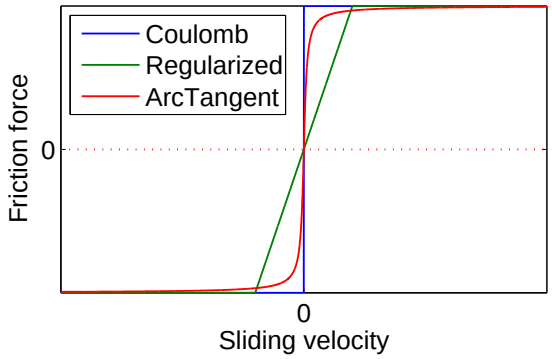


Figure 3.7: Sample Coulomb law regularizations

The basic regularization given by equation (3.7) is not perfect as its regularity is low, indeed, it is not differentiable at its transitions. An Arctangent law can be used to alleviate the problem,

$$\| \{f_T\} \| = \frac{2\mu \| \{f_N\} \|}{\pi} \arctan \left(k_t \| \{w\} \| \right) \tag{3.8}$$

Two main drawbacks exist with such regularization. First, parameter k_t does not have a clear physical meaning, so that its value is difficult to determine. In particular, when dealing with distributed contact on large surfaces, the contact pressure will be shown to have major variations over the surface. The use of a constant k_t may thus be a problem.

Second, the regularized contact laws loose their static capability. In (3.7), a case with no sliding velocity implies no friction force. This is not compatible with sticked states when friction force is needed to maintain the sliding velocity to zero. Only stick-slip transition can then be approached for a globally sliding configuration.

Regarding the thesis objectives – simulating large industrial models over long times – a simple friction law was chosen. A basic Coulomb regularization from equation (3.7) is then employed for the industrial applications. Using an Arctangent law did not provide different results for the simulations presented in chapters 4 and 5.

The introduction of richer contact laws is a clear perspective of this work. Although using an exact Coulomb law is possible, the time simulation application sought and presented in section 4.2 is not very compatible.

A last class of friction models is so-called *global friction* models and are presented in appendix D. Their aim is slightly different, as they typically provide an integrated friction behavior. Their application is then oriented towards the control in automatics involving systems with friction.

From a global point of view, all the effects of friction can be observed and analytical models reproducing each friction feature can be developed. Experimental identification can be performed rather easily as only the effect of friction on other measurable values is then of interest.

All the features provided by such analytical models should be present at a macroscopic scale when using simple local friction laws at the microscopic scale, presented above. The question is then to determine if the scale of the finite elements with a refinement suiting to linear mechanics of solids is relevant for friction. Caroli *et al.* [36, 37, 89] would rather state that it may not be the case. The use of global friction models at the scale of the finite elements could then be studied as a perspective of this work.

3.2.4 Numerical implementation

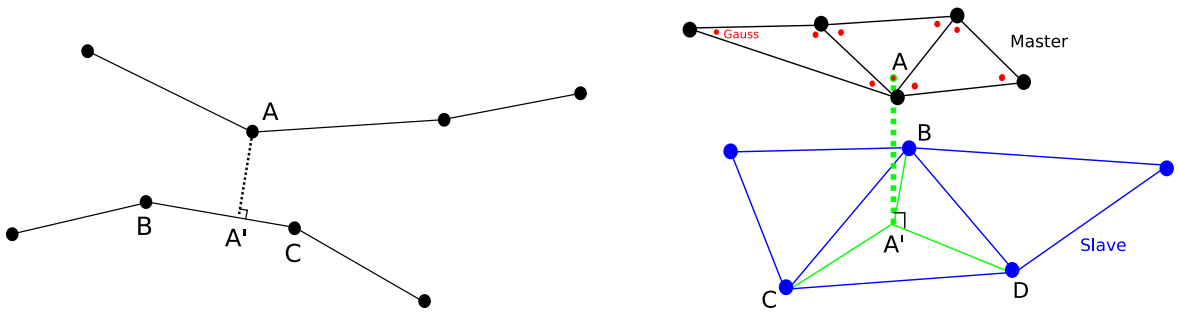
This section aims at presenting how the contact-friction models of section 3.2.1 can be applied to a 3D finite element model, and in particular how some adaptations have been carried out for a rotating-disc brake model.

The first feature to implement is the gap between two solids, as presented in figure 3.1. A slave/master strategy is commonly employed, so that the gap is evaluated by the projection of a contact point of the master surface onto the slave surface. These strategies are well documented, for example in the ABAQUS manuals [90] or code ASTER manuals [91].

Practically, the gap between two surfaces is defined as the relative displacement along the contact normal N with a possible offset g_0, as

$$\{g\} = N.(u_{slave} - u_{master}) - \{g_0\} \tag{3.9}$$

For a 2D application, as illustrated in figure 3.8a, the contact point A is projected as A' on the slave surface, defining the gap as the distance AA'. For a displacement u, the displacement of A' is expressed as a combination of the displacements of B and C, which are nodes of the slave surface model.



(a) 2D contact (b) 3D contact

Figure 3.8: Illustration of gap computation for finite element models

For a 3D application, as illustrated in figure 3.8b, the computations become more complex in implementation. This case is illustrated in more details as it corresponds to the actual implementation in SDT.

Once the master and slave surfaces are identified, the so called *contact points* must be chosen. They will be the points where the contact-friction forces will be actually computed for the mechanical equilibrium. In practice, they are points of the master surface, typically, the surface nodes are chosen, or in our case, the Gauss integration points of the underlying elements

(red dots in figure 3.8b. The interest in choosing Gauss points for contact resides in the fact that an element-wise handling is permitted which allows usual integration rules to be employed and simplifies handling of surfaces for the definition of contact pressures.

The identification of the projected A' of contact point A is referred as the *pairing* operation, as each contact point of the master surface must be paired to the slave surface depending on its configuration.

In practice, each contact point is first paired to an element of the slave surface, found as being the element in intersection with the outer normal of the contact point. The projected point A' is then expressed in position as a linear combination of the slave element points, B, C and D, which is commonly named *mapping*. Points B, C and D $\in \mathcal{N}_{\mathcal{S}e}$ the set of nodes of slave element e, they are computed using the shape functions of the slave element. They can be directly exploited, writing

$$u_{A'} = \sum_{x_k \in \mathcal{N}_{\mathcal{S}e}} \xi_k(A^{'}) u(x_k) \tag{3.10}$$

where $\xi_k(A^{'})$ are the coefficients relating the position of A' in e.

The gap defined by equation (3.9) can consequently be expressed for a contact point m_j as

$$g_j = N. \left(u(m_j) - \sum_{x_k \in \mathcal{N}_{\mathcal{S}e}} \xi_k(A^{'}) u(x_k) \right) - g_{j0} \tag{3.11}$$

In practice, the gap will be expressed at the contact points, but obtained from the mesh deformation u described by the DOF vector q. The practical equation to be exploited is then for a contact point j,

$$g_j = N. \left(\sum_{x_k \in \mathcal{N}_{\mathcal{M}e}} \left(\xi_k \left(A \right) q_k \right) - \sum_{x_k \in \mathcal{N}_{\mathcal{S}e}} \left(\xi_k(A^{'}) q_k \right) \right) - g_{j0} \tag{3.12}$$

where $\mathcal{N}_{\mathcal{M}e}$ (resp. $\mathcal{N}_{\mathcal{S}e}$) are the master element nodes containing (resp. the slave element nodes paired) with contact point j. $\xi_k(A)$ (resp. $\xi_k(A^{'})$ are the coefficients relating the position of contact point A in the master element (resp. of the projection $A^{'}$ of contact point A on the slave element). q_k is the displacement of node k.

Equation (3.12) is a linear combination of the system DOF, so that it can be summarized using an observation matrix $[C_{NOR}]$. The gap vector expressed at each contact point is then written

$$\{g\}_{N_c \times 1} = [C_{NOR}]_{N_c \times N} \{q(u)\}_{N \times 1} - \{g_0\}_{N_c \times 1} \tag{3.13}$$

noting N the number of system DOF, and N_c the number of contact points.

The force resulting from the gap-pressure relationship, with the gap defined at each Gauss point by equation (3.13), is then defined as

$$\{\hat{q}\}^T \{f_N\} = \int_\Gamma \{\hat{u}(\hat{q})\}^T N p dS \simeq \sum_e \sum_j \{\hat{u}(\hat{q})\}^T \{N\} p(x_j, q) \omega_j^{(e)} J^{(e)}(x_j) \tag{3.14}$$

where f_N is the global contact force, p the contact pressure, $\hat{q}$ a virtual displacement, q the displacement, $x_j^{(e)}$ are the integration points of current element e, $J^{(e)}(x_j)$ the Jacobian of the shape transformation (surface associated to each integration point) and $\omega_j^{(e)}$ the weighting associated with the integration rule of element e.

The use of the observation matrix $[C_{NOR}]$ is also possible, so that in practice the contact force expressed at the contact points can be recovered at the nodes using

$$\{f_N\}_{N \times 1} = [C_{NOR}]^T_{N \times N_c} \left\{ \omega_j^{(e)} J^{(e)}(x_j) \; p(x_j, q) \right\}_{N_c \times 1} \tag{3.15}$$

The computation of the sliding velocity requires the computation of the differential velocities of the slave and master surfaces,

$$\{w\} = T.(\dot{u}_{slave} - \dot{u}_{master}) \tag{3.16}$$

where T represents the directions in the friction plane, defined as orthogonal to the contact normal N. T can consequently represent two directions, and the sliding velocity vector be of size $2N_c$. The same developments than for the computation of the gap can be performed, leading to the generation of a tangential displacement observation matrix $[C_{TAN}]$, from which the sliding velocity is

$$\{w\}_{2N_c \times 1} = [C_{TAN}]_{2N_c \times N} \{\dot{q}\}_{N \times 1} \tag{3.17}$$

and the friction forces can be recovered, from the local friction force f_{tj}

$$\{f_T\}_{N \times 1} = [C_{TAN}]^T_{N \times 2N_c} \left\{ \omega_j^{(e)} J^{(e)}(x_j) \; f_{tj}(p(x_j, q)) \right\}_{2N_c \times 1} \tag{3.18}$$

For disc brake squeal applications, the pad/disc contact surface has a specific topology, which is presented in figure 3.9. To have a constant sign for disc velocity, the positive *tangential* direction e_θ is taken identical on the upper and lower pads. The outer normals e_z however differ on the upper and bottom sides, so that the radial direction e_r is towards the exterior on the lower (*inner*) pad and the interior on the upper (*outer*) one.

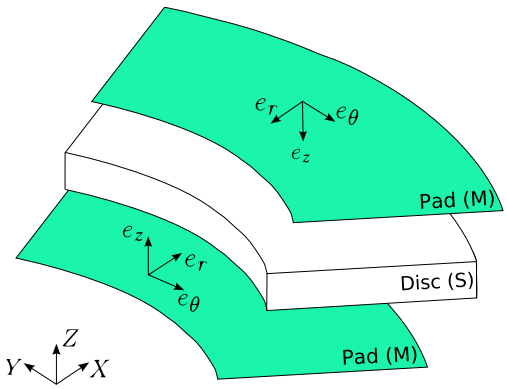


Figure 3.9: Local frame definition of the pad/disc contact

Local frame handling is performed when building the observation matrices (3.13) and (3.17). If the disc were actually rotating, the pairing operation previously described should be performed every time the mesh displacement is updated. It is an intensive operation when handling numerous contact points of an industrial model (here around 1,800 contact points), especially if it must be performed at each time step of a transient simulation of over 100,000 steps.

Considering a rotating local frame based on the frame presented in figure 3.9 is then a solution to compute pairing once for all. The effect of the disc rotation must then be directly taken into account in the system mechanical equations, which will be referred as an *Eulerian* formulation. As proposed by Moirot [5], the disc mesh will remain fixed while the disc material rotates inside. Numerically, the continuous displacement u of the disc can be expressed in the rotating frame as

$$u = u(r, \theta(t), z, t) \tag{3.19}$$

The frame is considered to rotate at a constant velocity Ω, around the e_z direction. As the time span possible from the transient simulations provided in chapter 4 is in the order of the tens of milliseconds the effect of the disc slowing down is neglected, so that

$$\theta(t) = \theta + \Omega t \tag{3.20}$$

The velocity of the disc can then be derived from equation (3.19) considering the total derivative of u, which implies the variations of u as function of its time dependant parameters and the rotation of the base vectors,

$$v = \frac{du}{dt} = \frac{\partial u}{\partial t} + \frac{\partial u}{\partial \theta}\frac{\partial \theta}{\partial t} + \Omega e_z \wedge u \tag{3.21}$$

which can be simplified using (3.20),

$$v = \frac{\partial u}{\partial t} + \Omega \frac{\partial u}{\partial \theta} + \Omega e_z \wedge u \tag{3.22}$$

The sliding velocity is the difference between the disc velocity computed in the rotating frame in equation (3.22) and the velocity of the pad defined in the fixed frame, so that

$$w = v_{pad} - v_{disc} = \frac{\partial u_{pad}}{\partial t} - \frac{\partial u}{\partial t} - \Omega \frac{\partial u}{\partial \theta} - \Omega e_z \wedge u \tag{3.23}$$

The disc rotation provides a natural sliding direction. It was thus chosen in this work to ignore radial velocities. The projection of equation (3.23) on e_θ yields

$$w_\theta = (v_{pad} - v_{disc})\, e_\theta - \left(\Omega \frac{\partial u}{\partial \theta}\right) e_\theta + \Omega u \cdot e_r \tag{3.24}$$

where the product $(e_z \wedge u) \cdot e_\theta$ from the projection of equation (3.23) on e_θ is simplified as the projection of the positions on the radial direction.

In (3.24), the sliding velocity can be decomposed in three contributions from left to right a vibration, convection, and rotation terms can be observed. The first one comes from the fixed mesh vibrations, which are actually computed on the system. The rotation term is constant and only depends on the mesh original topology. The convection term is more difficult to implement as it implies spatial derivations in cylindrical coordinates with a mesh in Cartesian coordinates. This term is thus ignored in [6].

Within an element, displacement u is related to degrees of freedom q through shape functions N

$$u = Nq \tag{3.25}$$

the convection term is derived as

$$\frac{\partial u}{\partial \theta} \cdot e_\theta = \frac{\partial N}{\partial x_j}(q \cdot e_\theta)\frac{\partial x_j}{\partial \theta} \tag{3.26}$$

where x_j represents the Cartesian directions.

In the contact implementation used for this work, shape function utilities of SDT are used to generate an observation matrix $[C_{N\theta}]$ relating convective contributions of the sliding velocity to the vector of DOF q as in equation (3.26).

In practice, the sliding velocity is thus computed using

$$\{w_\theta\}_{N_c \times 1} = [C_{TAN}]_{N_c \times N}\, \{\dot{q}\}_{N \times 1} + \Omega\, [C_{N\theta}]_{N_c \times 1}\, \{q\}_{N \times 1} + \Omega\, \{r\}_{N_c \times 1} \tag{3.27}$$

where the tangent displacement observation matrix $[C_{TAN}]$ has been projected on the tangential direction e_θ, and $\{r\}$ is the radius of each contact point to the rotating frame origin.

The effect of the convection term is interesting to interpret, it accounts for the fact that the disc has a static *Eulerian* deformation due to friction. The disc slows down entering the pad area and accelerates when existing. The effect is obviously a linear function of the friction coefficient.

Figure 3.10 illustrates the convection term expression on a simplified disc section mesh (from the simplified brake model). For a given element, the convection contribution is a measure of the element compression along the tangential direction illustrated by the $[C_{N\theta}]$ weights illustrated in figure 3.10.

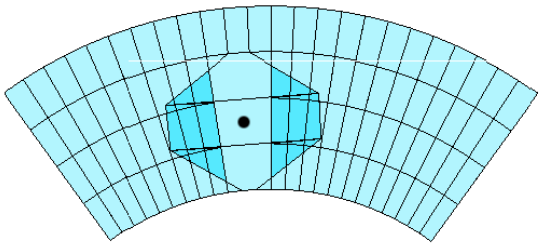


Figure 3.10: Weight for the convection terms at point (•)

In application to realistic brakes, the convection terms are plotted for the 12 Bar static state with a disc rotation of $5rad/s$ in figure 3.11. The regularity is not perfect, this is mostly due to the fact that the static state has been computed in ABAQUS, and post-treated in SDT. The regularization that ABAQUS would use may not have been applied. The values are rather small, clear trends can however be observed.

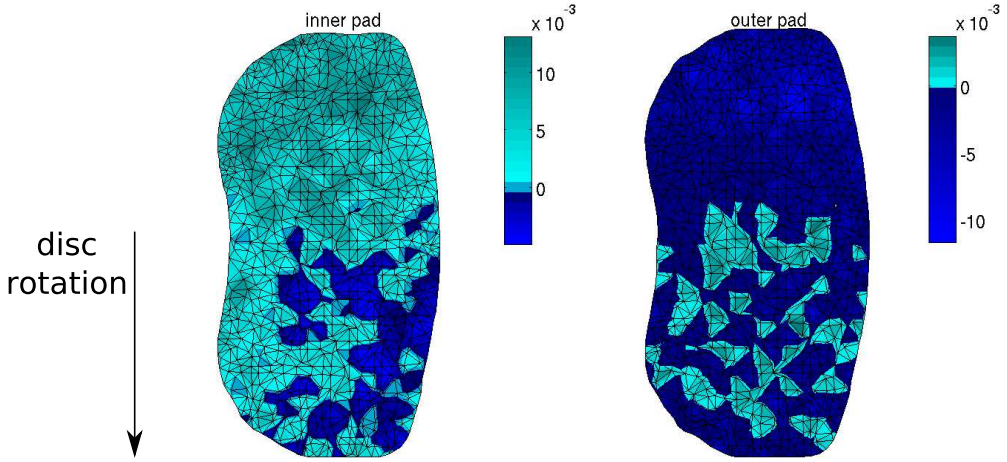


Figure 3.11: Convection term $\frac{\partial u_\theta}{\partial \theta}$ [mm] for the real brake pads at 12 Bar. Data plotted on a Delaunay triangulation of the Gauss contact integration points

In figure 3.11, the disc is rotating from top to bottom. The outer pad presents a clear compression at the entry (negative convection term) while some extension is seen at the exit. The outer pad load is indeed applied by fingers at both pad sides. The loading of the inner pad is very different as the load is located under the center (piston position), so that the disc is not loaded at the entry, a clear extension is then observed. Looking at both convections, it is clearly seen that the disc is deformed at the entry with a shear force due to friction and the asymmetric pad loading.

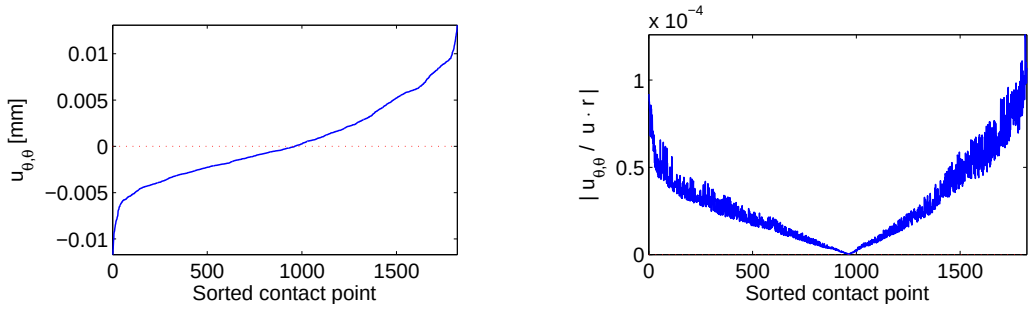


Figure 3.12: Convection term $J(x_i)\frac{\partial u_\theta}{\partial \theta}_i$ [mm] sorted in increasing order (left) for both pads and comparison (right) to the disc rigid rotation term at 12 Bar, for a disc rotation velocity of $5 rad/s$

Figure 3.12 plots the sorted convection terms which globally shows the disc entry/compression, exit/extension pattern. The convection terms are however much smaller than the disc rigid rotation terms. As plotted in figure 3.12 the rigid rotation is 10,000 times higher on average. The convection feature, although implemented for the simulations, may consequently not be of a particular interest. Its effect on the results is indeed negligible.

The disc rotation also has an effect on the expression of the acceleration for the formulation of the mechanical equilibrium. The acceleration of the disc is defined as

$$a = \frac{dv}{dt} \tag{3.28}$$

which is obtained by deriving equation (3.22),

$$a = \frac{\partial^2 u}{\partial t^2} + 2\Omega \left(\frac{\partial^2 u}{\partial t \partial \theta} + e_z \wedge \frac{\partial u}{\partial t} \right) + \Omega^2 \left(\frac{\partial^2 u}{\partial \theta^2} + 2 e_z \wedge \frac{\partial u}{\partial \theta} \right) - \Omega^2 \left(u \cdot r \right) \tag{3.29}$$

where r is the projection vector on the tangent plane $\{e_r; e_\theta; 0\}$.

Equation (3.29) presents several additional terms to the displacement acceleration $\frac{\partial^2 u}{\partial t^2}$. From left to right, a gyroscopic damping associated to the first order time derivatives of u (2Ω factor), a gyroscopic stiffness associated to the θ only derivatives of u (Ω^2 factor) and a volume centrifugal load are identified.

The mechanical equilibrium equation should then be altered to stand for all rotational effects in the disc. In practice however, the disc rotation velocities targeted are rather small – a few rad/s – so that all gyroscopic and centrifugal terms are neglected, for the applications presented in chapters 4 and 5, thus leading to

$$a \simeq \frac{\partial^2 u}{\partial t^2} \tag{3.30}$$

Adjusting rotation related terms in the equilibrium equation is a perspective for future work.

The numerical implementation provided here through the gap equation (3.13), exponential pressure-gap relationship from table 3.1, contact force formulation (3.15), sliding velocity (3.27), friction model equation (3.7) and friction force formulation (3.17) is sufficient to perform time simulations as presented in section 4.2. Given a displacement-velocity state expressed on the system DOF, contact-friction forces can be computed and added as non-linear forces to the equilibrium equations of the Newmark time integration scheme presented in section 4.1.4.

3.2.5 Frequency domain linearization

When characterizing a non-linear system in the frequency domain, the most common approach is to consider the system linearized around a deformation state and compute the modes of the linearized system. The linearized model is then called *tangent* to the said deformation state. This strategy is of course limited as linearizing in the vicinity of deformation states can only be relevant for very small deformations from the targeted working point.

The expression of the system linearization strongly depends on the chosen contact formulation. For Lagrangian formulations, the definition of the tangent state only depends on the surface effectively in contact, as a contact point is either *open* (no contact) or *closed* (in contact). The open case generates no contact stiffness, while the closed case generates an exact continuity constraint for the normal displacement at the interface.

Penalized contact formulations are defined as function of the gap, such that a tangent state can be derived for each contact point by deriving the gap/pressure relation as illustrated in figure 3.13. It should be noted that the contact stiffness differs at each contact point and depends on the local gap.

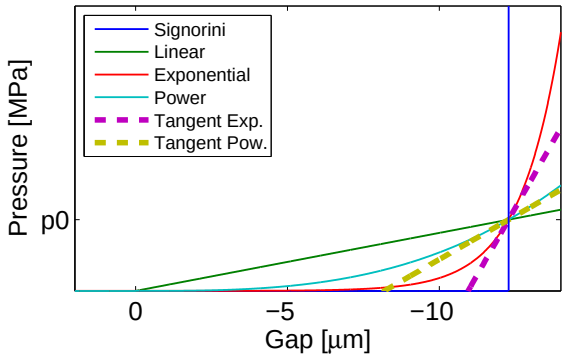


Figure 3.13: Illustration of tangent contact states at a given pressure state

For the exponential contact law used here, at a given contact point i the contact stiffness k_{ci} is defined by the derivative of exponential law equation of table 3.1

$$k_{ci}(q) = \omega_i J(x_i) \frac{\partial p_i(q)}{\partial q} = -\omega_i J(x_i) p_0 \lambda e^{-\lambda g_i} \tag{3.31}$$

where g_i is the gap at a contact point i observed at the i^{th} line of the gap observation matrix $[C_{NOR}]$, $J(x_i)$ the corresponding surface, and w_i the Gauss weighting coefficient. Using the gap observation equation (3.13) and the force restitution of equation (3.15), the tangent contact stiffness matrix $K_{nlc}(q)$ is recovered,

$$K_{nlc}(q) = [C_{NOR}]^T \left[\backslash w_i J(x_i) k_{ci}(q) \backslash \right] [C_{NOR}] \tag{3.32}$$

K_{nlc} is therefore symmetric positive definite and can be used for real and complex modes.

The effect on the system itself is a perceived stiffness variation occurring at each point of the contacting interfaces. The distribution of contact stiffness can then show large variations over the surface, as will be illustrated in chapter 5.

The definition of a tangent friction state is more difficult, as it couples normal and tangential directions, it is typically non-symmetric – it is after all the only instability source for the squeal simulations.

The nature of friction as written in equation (3.2) shows that a variation of contact force has a direct effect on friction force, while a variation of the friction force may occur without effect on the contact force. This non symmetric notion is illustrated in figure 3.14 where stiffness coupling matrix topologies are illustrated for a basic case.

In the case of a sliding state, the friction force is explicitly defined and only depends on the contact force at the same point. The tangent friction is then the tangent contact (3.31) scaled by the friction coefficient μ, linking the normal and tangent displacements for complex modes. For real mode computation, since the tangent displacement is free for a fixed contact state, no stiffness is added. The tangent friction coupling stiffness K_{nlf}^{slide} for sliding states is defined as

$$K_{nlf}^{slide} = \begin{cases} 0 & \text{for real mode computation} \\ [C_{TAN}]^T \left[\backslash \mu k_{ci}(q) \backslash \right] [C_{NOR}] & \text{for complex mode computation} \end{cases} \tag{3.33}$$

The formulation of tangent friction matrices of equation (3.33) is very classical, and is commonly employed under the *small sliding perturbations* formulation. This was used by Moirot [5] to Lorang [6] for train brake squeal applications, or by Vola *et al.* [92] to study rubber/glass instabilities in sliding steady states. It was also studied more academically by Nguyen [69], Pires and Trabucho [93], and Andersson and Klarbring [94].

For 3D solids, Moirot [5] noted that the sliding friction tangent state also features a damping term due to the possibility for the friction force to change its orientation. Indeed, the sliding velocity can change its direction due to the vibration, since it is located in a plane. For an arbitrary variation of the sliding velocity $\delta\dot{q}$, the friction force writes

$$f_T(\dot{q} + \delta\dot{q}) = \frac{1}{\|w\|} \left(\mu f_N \left[C_{TAN} \right]_\theta (\dot{q} + \delta\dot{q}) \right) + \mu f_N \left[C_{TAN} \right]_r (\delta\dot{q}) \tag{3.34}$$

A variation of the velocity in e_θ does not induce a variation of f_T. A friction force variation in the radial direction e_r induces a change in f_T, given by

$$f_T(\dot{q} + \delta\dot{q}) - f_T(\dot{q}) = \frac{1}{\|w\|} \mu f_N \left[C_{TAN} \right]_r \delta\dot{q} \tag{3.35}$$

Considering the system tangent state, equation (3.35) thus generates an additional damping term, which opposes the variation of the sliding direction towards the radial direction,

$$\delta f_T = -\left[C_{fr} \right] \delta\dot{q} = -\left[C_{TAN} \right]_r^T \left[\backslash \frac{1}{\|w_i\|} \mu \left(f_{Ni} \cdot e_z \right) \backslash \right] \left[C_{TAN} \right]_r \delta\dot{q} \tag{3.36}$$

Although a minus term appears in equation (3.36) the actual effect is positive. Indeed, the projection of the contact force on the local frame of figure 3.9, $(f_{Ni} \cdot e_z)$ generates a negative term. e_z is the outer normal of the master surface, and the contact force is compressive *i.e.* oriented towards the inner normal.

In the case where penalized contact laws are used, the contact force can become a traction force and then generate some additional instability. The contact force amplitude in a traction configuration has however to be negligible to be physical so that no problem arises in practice.

The PhD of Moirot [5] provides several results in terms of existence and unicity of solutions in the sliding perturbations formulations. Considering the friction states do not change for the modal computation, it allows the application of Lyapunov theorem.

For contacting interfaces other than the pad/disc, the static friction state is sticking. The basic Lagrangian approach will consider an exact continuity constraint of the tangential displacement at the interface, thus making it a bilateral constraint. Taking such pattern into account

is critical for the complex modes as a physical resistance to the tangential displacement occurs. For sliding to happen, the friction force must increase up to the sliding threshold, this friction force variation must then be reproduced.

A penalized approach can be used, which relaxes the Lagrange constraint. The friction force is however unknown inside the Coulomb cone, and further studies are necessary to characterize the stiffness coupling levels and the system sensitivity. Experiments are suggested by Bureau *et al.* [89, 36] for non metallic materials.

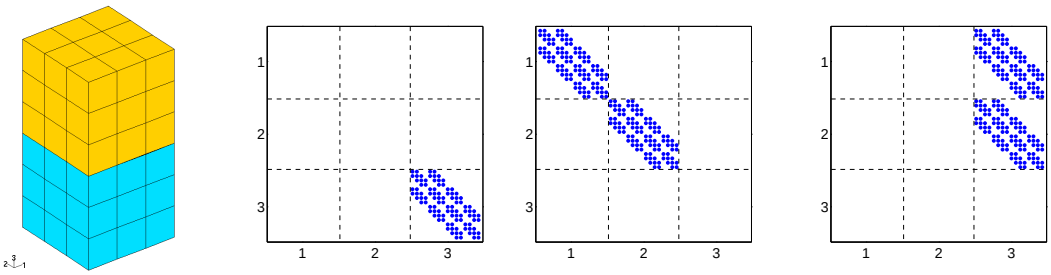
The strategy chosen here is to consider the Coulomb cone limit, characterized by the contact force level, illustrated in figure 3.2. A proportional relationship linking the tangent friction stiffness level and contact levels is set, using parameter ν. The stiffness behavior will thus be close to microslip patterns, occurring for low vibration levels. Obviously ν, which has to be somehow identified, will verify

$$0 \leq \nu \leq \mu_s \tag{3.37}$$

The tangent friction coupling stiffness K_{nlf}^{stick} for sticking states is then defined for both real and complex mode computation as

$$K_{nlf}^{stick}(q) = [C_{TAN}]^T \left[\backslash \nu w_i J(x_i) k_{ci}(q) \backslash \right] [C_{TAN}] \tag{3.38}$$

To illustrate the effect of the different tangent contact-friction matrices presented here, an example based on a very simple problem is presented. The basic model given in figure 3.14a features two cubes in vertical contact. For illustration purposes, the degrees of freedom are segregated by their direction, 1 and 2 define the horizontal directions while 3 is the vertical normal direction, in which contact is defined.



(a) Cube model (b) Contact (c) Sticking friction (d) Sliding friction

Figure 3.14: Tangent contact-friction matrix topologies on a simple case. Direction 1,2 are horizontal (friction plane). Direction 3 is vertical (contact direction).

Figure 3.14 presents the topologies of the tangent contact-friction matrices expressed on the master surface DOF. The tangent contact matrix (3.32) couples the normal directions, here the third one, as shown in figure 3.14b. The matrix is symmetric positive definite and used for real or complex mode computation.

The tangent friction matrix (3.38) for sticking states is plotted in figure 3.14c. It couples directions 1 and 2 independently, is symmetric positive definite and is used in real and complex modes computation.

For sliding state, the tangent friction matrix (3.33) is shown in figure 3.14d. The unsymmetric coupling is clearly observed as direction 3 is coupled to both directions 1 and 2. A deformation along 3 has an effect on directions 1 and 2 but not the other way around.

3.3 Classical mechanical instability analyses

From a non-linear system, stability is usually computed from a *static* or *pseudo-periodic* equilibrium solution, using classical Lyapunov stability concepts presented in section 3.3.1. Such analyses are however limited to the initial divergence and do not provide information about vibration levels achieved later on. Sections 3.3.2 and 3.3.3 provides a rapid review of the very rich literature on methods for the computation of limit cycles and the related non-linear normal modes. This review provides useful insight on how non-linear responses are analyzed. The main constraint of this work was the imperative to run computations on an industrial model of a full brake. Within this constraint, a number of analyzes were possible and the choices made are listed in section 3.3.4.

3.3.1 Stability computations

The first operation performed to assess instabilities of a system with non-linearity, is to consider its propensity to diverge from equilibrium. In the classical characterization of non-linear instabilities, the system equations are written using a non-linear state-space formulation, with state vector q,

$$\{\dot{q}(t)\} = A(\{q(t)\}, \{\nu\}) \tag{3.39}$$

where A is a non-linear application function of states and so-called *control parameters* $\{\nu\}$. Classical developments in the literature only consider a set of global parameters $\{\nu\}$ such as the friction coefficient. A novelty of this thesis, through the CMT method presented in section 2.3, is to add component modes as control parameters.

The state vector $q(t)$ origin is typically set on an equilibrium solution of the system, so that with a proper offset, $q = 0$ is a solution of equation (3.39). The study is then focused on the system behavior in the case where small deviations from the equilibrium occurs, *i.e.* if instabilities occur when the system is externally excited, for example by a variation of a control parameter ν_j.

Under regularity conditions, the Taylor expansion of equation (3.39) gives an approximation of the system behavior. Considering a first order expansion, or *linearization*, one has

$$\dot{q}(t) \simeq \frac{\partial A}{\partial \delta q}(q(t), \nu) \tag{3.40}$$

Equation (3.40) is commonly referred to the system tangent to the non-linear equations at a given equilibrium point. The linearized system can be written in matrix form

$$\dot{q}(t) = [A(\nu)]\, q(t) \tag{3.41}$$

The study of the system modes, eigenvalues and eigenvectors of $[A(\nu)]$ (see appendices A, and B), gives the approached response of the non linear system in the vicinity of the equilibrium. The free response of system (3.41) can indeed be expressed as

$$\{q(t)\} = \sum_j \left(\{a_j\}\, e^{\lambda_j(\nu)t} + \{\bar{a}_j\}\, e^{\bar{\lambda}_j(\nu)t} \right) \tag{3.42}$$

with λ_j the eigenvalues of $[A(\nu)]$.

From equation (3.42), the concept of stability is obvious. Depending on the eigenvalues of $[A(\nu)]$, the response will be stable, or unstable. Under the formalism of Lyapunov [95], the stability of an equilibrium q_e is characterized by the fact that for a given instant t_0, if the system state is close to q_e, it will converge back to q_e, in more mathematical terms,

$$\forall \epsilon > 0, \exists \delta > 0 / \|q(t_0) - q_e\| < \delta \Rightarrow \forall t > t_0, \|q(t) - q_e\| < \epsilon \tag{3.43}$$

The characterization given in (3.43), is the classical asymptotic Lyapunov stability. Other definitions by Poincaré for generalization purposes (orbital stability) are also well known [96].

Combining the tangent response of the linearized system of equation (3.42) and the stability conditions (3.43) leads to the classical result that the system is stable if the real parts of the eigenvalues of $[A(\nu)]$ are negative or zero. If at least one eigenvalue has a positive real part, divergence will occur and one can expect that with amplitude growth, the approximation (3.40), will no longer be valid.

For static and pseudo-static applications (case of rotating frames), the application of Lyapunov theorem to equation (3.42) is a sufficient condition to detect instabilities. One thus classically analyzes the conditions on control parameters $\{\nu\}$ for which the linearized system is unstable. Most work on brake squeal use at least this approach. In the present work, the CMT method is used to study the impact of changes in the component behavior, section 2.3. More classical studies on friction and pressure parameters are given in appendix C.

For systems with periodic responses, the same notions can be exploited, using the Floquet theory. The stability is then evaluated using a *monodromy* matrix. State-of-the-art computations are for example proposed by Lazarus *et al.* [97]. One can also note that alternate *equivalent linearization* procedures exist, see Iwan [98] for example.

Following the stability evolution as function of the parameters, a common phenomenon called *lock-in* can happen, where two previously complex-conjugate eigenvalues of (3.42) become opposite real, thus generating instability. This phenomenon is well documented in many domains, *e.g.* by de Langre [99] in fluid structure interactions. It has also been experimentally observed for brake squeal applications by Massi *et al.* [7, 100]. The friction coefficient is then commonly considered as a varying parameter. This can be discussed nowadays, as it was shown numerically by Linck [8] and verified in the thesis context [21] that the variation of the global friction coefficient can also be the consequence of the system vibrations, although using a local constant friction coefficient.

The transition of a system from a stable state to another state (stable or not) is commonly called *bifurcation* as first introduced by Poincaré. Numerical estimations of such bifurcation are possible, using the global parametric approaches aforementioned. A complete zoology of bifurcation types has been developed historically depending on the way the eigenvalues evolve [96]. One of the most common bifurcation type and which is of interest here is the Hopf bifurcation. This happens for mode lock-in. The characterization of a Hopf bifurcation point states that if an eigenvalue transits from a negative real part to a null one, with a non null velocity for a specific value ν_c of parameter ν,

$$\begin{cases} \Re\left(\lambda_j(\nu_c)\right) = 0 \\ \left.\frac{d\lambda_j(\nu)}{d\nu}\right|_{\nu=\nu_c} \neq 0 \end{cases} \tag{3.44}$$

and if the conditions are verified for the existence of a periodic movement of constant amplitude – a *limit-cycle* – the system will transit to it, following the classical scheme of figure 3.15.

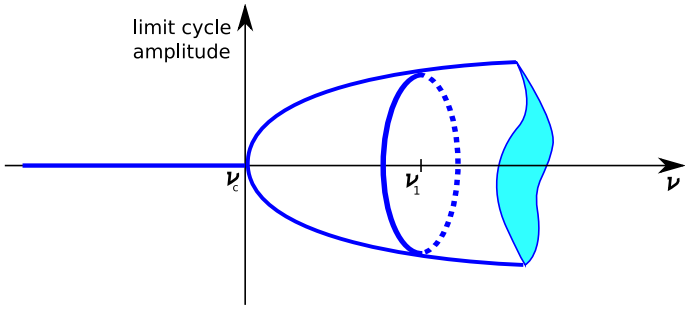


Figure 3.15: Illustration of a Hopf bifurcation point

3.3.2 Limit cycle computations

Although the stability computations give indications on the behavior of a non-linear system in the vicinity of the equilibrium, the validity range is unknown and can be potentially small. Furthermore, the analysis does not provide the final amplitude of the system after the instability started to express. If several non linearities are present, their combination can also yield non trivial behavior that must be assessed by a non linear study.

Assuming the existence of limit cycles predicted by the Hopf bifurcation theory, a number of methods exist for the estimation of the non-linear response. The *center manifold*, *multiple scales*, *normal forms*, and *harmonic balance* methods will be summarized here as they can be combined for the desired prediction. Although some of these methods have been deployed to relatively complex systems in terms of DOF number, they are often too complex to be applied to industrial models.

The center manifold method has been used typically as a preliminary method since it reduces the system to consider, which is of interest for larger models. It is well described in the book of Carr [101], and has been used for example by Pesheck *et al.* [102] to compute non-linear normal modes, or by Sinou *et al.* for aircraft brake applications [67] or automotive brake squeal applications [103].

Considering the linearized system (3.41), characterized by its modes, one separates modes with a negative, positive and null real part, thus generating three subspaces, from which tangent manifolds exist. In practice, as the study is performed from a stable state, the unstable manifold is supposed empty, and only the stable and centered subspaces remain.

The displacement vector can be expressed as a linear combination of the system modes (generalized coordinates), from which a DOF segregation is performed between stable DOF q_s and centered DOF q_c. Writing the displacement vector $q = \left\{ \begin{array}{c} q_s \\ q_c \end{array} \right\}$, equation (3.39) can be split in two and expressed as the addition of a linear part and a non-linear part,

$$\left\{ \begin{array}{l} \dot{q}_s = [A_s]\, q_s + \mathcal{F}(q_s, q_c) \\ \dot{q}_c = [A_c]\, q_c + \mathcal{G}(q_s, q_c) \end{array} \right. \tag{3.45}$$

The center manifold theory states then that the stable manifold is invariant, and that the centered one can vary. The stable manifold can thus be expressed as function of the centered one,

$$q_s = h(q_c) \tag{3.46}$$

where h verifies the tangent property of the manifold, namely $h(0) = 0$ and $\frac{dh}{dq_c} = 0$. Using equation (3.45) and (3.46), and the fact that $\dot{q}_s = \dot{q}_c \frac{dh}{dq_c}$,

$$\frac{dh}{dq_c} \left([A_c]\, q_c + \mathcal{G}(q_s, q_c) \right) = [A_s]\, h(q_c) + \mathcal{F}(h(q_c), q_c) \tag{3.47}$$

Using polynomial expansions for the non-linear terms of $\mathcal{F}$ and $\mathcal{G}$ and h allows the identification of the coefficients of h expansion in equation (3.47). The initial system is then reduced to its center manifold,

$$\dot{q}_c = [A_c]\, q_c + \mathcal{G}(h(q_c), q_c) \tag{3.48}$$

The sets of parameters ν and their value of interest ν_0 verifying $\nu = \nu_0 + \epsilon$, can then be introduced, writing

$$\left\{ \begin{array}{l} \dot{q}_c = [A_c(\nu)]\, q_c + \mathcal{G}(h(q_c), q_c, \nu_0) \\ \dot{\nu} = 0 \end{array} \right. \tag{3.49}$$

where the trivial second equation is necessary to the resolution.

Equation (3.49) then describes the initial system on its unstable (or rather not stable) modes thus considerably reducing its size. Depending on the non linearity, the expansion order of h may reach several orders, which can become cumbersome. This approach is commonly used however to re-cast an original system as a starting point for another non-linear method.

The *multiple scales* method was for example employed by Nayfeh *et al.* to compute non linear normal modes [104]. The philosophy is here to consider the system state evolution independently from the non-linear evolution, introducing *quick* time scales for the system oscillation and *slow* time scales for the non-linearity evolutions. The solution is then written

$$q(t,\epsilon) = q_0(T_0, T_1, T_2, \cdots) + \epsilon q_1(T_0, T_1, T_2, \cdots) + \epsilon^2 q_2((T_0, T_1, T_2, \cdots) + \cdots \tag{3.50}$$

where $T_0 = t$, $T_1 = \epsilon t$, $T_2 = \epsilon^2 t$. The differentiation of equation (3.50) gives a development of the solution q, which can be resolved order by order, from 0 to n. The constants found for each order, are then considered as dependent on the lower scale. Typically, implementing the solution form (3.50) in the system equation (3.39), the order zero equation is the linear response, from which q_0 can be extracted, using the eigenvalues of $\frac{dA}{dT_0}$

$$q_0(t) = \sum \alpha_k(T_1) e^{\lambda_{0j} t} \tag{3.51}$$

the first order equation is then considered in which the solution q_0 is cast to solve q_1, and so on.

The *normal forms* method starts from the center manifold reduction, and has been used for example by Hsu [105, 106] for the analysis of critical and post critical behaviors, and Jezequel *et al.* [107]. The aim of the normal form method is to express the Taylor expansion of the center manifold equation (3.49) in a better way.

The Taylor expansion is written as

$$\dot{q}_c = [A_c]\, q_c + \sum_{k=1}^{n} \frac{d^k \mathcal{G}}{dt^k}(q_c) + \mathcal{O}(q_c^{n+1}) \tag{3.52}$$

The normal form theory (Poincaré-Dulac) then states that a sequence of transformations $q_c = u + \mathcal{P}(u)$ where $\mathcal{P}$ is a vector of homogenous polynoms, allows to re-cast the system,

$$\dot{u} = [A_c]\, u + \sum_{k=1}^{n} g^{(k)}(q_c) + \mathcal{O}(q_c^{n+1}) \tag{3.53}$$

where $g^{(k)}$ have specific properties. The space of $g^{(k)}$ is orthogonal to the image of the endomorphism $[A_c]\, q_c$ applied to h. Equation (3.53) is the normal form, from which, if the eigenvalues of $[A_c]$ are not resonant, the system can be re-cast into a linear one. In the other case, the normal form can be identified for simulations. The number of terms kept $g^{(k)}$ is minimal and the final formulation is optimized.

The *Harmonic Balance Method* (HBM) is one of the most widespread non-linear method with industrial applications, due to its implementation easiness and good results. It is typically employed for forced response computations, or used when the period is unknown. It was used by Petrov and Ewins for turbomachinery applications, modeling contact at blade roots interfaces in [108, 109, 110], or VonGroll and Ewins for rotor stator problems [111]. Guskhov *et al.* proposed an application to squeeze film bearings [112] using a multi-dimensional HBM. An iterative implementation was proposed by Pierre *et al.* [113] for friction studies in damped systems. Recent applications by Coudeyras [114, 115, 11] and Jaumouillé [116] featured generalizations of the HBM. Coudeyras focused on the multi-dimensional HBM for autonomous systems using a *constrained* form of the HBM. Jaumouillé [116] presented an adaptative formulation depending on the level of non-linearity in the system as function of the frequency.

Only the concept of HBM will be presented here, most recent applications featuring a much greater degree of refinement. The HBM approximates the non-linear solution as a development truncated at order m of harmonics of an unstable frequency ω,

$$q(t) = a_0 + \sum_{k=1}^{m} \left(a_k \cos(k\omega t) + b_k \sin(k\omega t) \right) \tag{3.54}$$

The displacement is thus described by the Fourier coefficients a_k, b_k. Equation (3.54) can be written using the a vector of Fourier coefficients $z = \{a_0 a_1 b_1 \cdots a_N b_N\}^T$, and a time dependent matrix of trigonometric functions $T(t)$,

$$\{q(t)\}_{N\times 1} = [T(t)]_{N\times N(2m+1)} \{z\}_{N(2m+1)\times 1} \tag{3.55}$$

Applying classical mechanical equations, displacement and velocities are derived from equation (3.54), yielding

$$\begin{aligned} &\sum_{k=1}^{m} \Big(\Big([K] - (k\omega)^2 [M] \Big) a_k + k\omega [C] b_k) \Big) \cos(k\omega t) + \\ &\quad \sum_{k=1}^{m} \Big(\Big([K] - (k\omega)^2 [M] \Big) b_k + k\omega [C] a_k) \Big) \sin(k\omega t) \quad = \{f_c\} + \{f_{NL}(a_k, b_k)\} \end{aligned} \tag{3.56}$$

A difficulty arises if the non-linear forces cannot be expressed easily by a Fourier series. In such case, an iterative algorithm must be introduced. For a given state of the Fourier coefficients, the displacement is recovered in the time domain by an inverse FFT and the corresponding non-linear forces are computed in the time domain for each instant. The Fourier coefficients of the non-linear force are then obtained by the FFT of its transient expression.

Equation (3.56) is resolved by Galerkin identification, using the orthogonality of the trigonometric functions. This yields $m+1$ independent systems of 2 equations,

$$[\Lambda] \{z\} = \{f_c\} + \{f_{NL}(z)\} \tag{3.57}$$

where $[\Lambda]_{N(2m+1)\times N(2m+1)} = \left[\backslash \Lambda_k \backslash \right]$, is block diagonal, with $[\Lambda]_0 = [K]$ and

$$[\Lambda]_{k\geq 1} = \begin{bmatrix} [K] - (k\omega)^2 [M] & k\omega [C] \\ -k\omega [C] & [K] - (k\omega)^2 [M] \end{bmatrix} \tag{3.58}$$

Since the system size is now of size $N(2m+1)$ instead of N, its resolution can be computationally very demanding. Typical system reduction in the case of local non-linearity at this stage involves reduction to the non-linear DOF. If the number of non-linear DOF is *small enough i.e.* under a few hundreds, the linear DOF can be condensed on the non-linear ones, and the whole resolution is performed on the condensed form of equation (3.57).

3.3.3 Non linear normal modes

The most straight forward approach to compute limit cycles is in fact the use of full transient analyses, as proposed in the present thesis, and which was for example recently used by Lorang [6], Sinou [117], Ouyang [13], Linck [8] or Meziane [9], in the domain of brake squeal.

The application of time simulations is generally limited, in terms of computational cost and exploitation. The developments presented in sections 4.2 and 4.4 suggest new methods to alleviate such limitations. Time simulations are however commonly avoided in the literature. A lot of applications thus target other ways of characterizing a non-linear system, using generalizations of the canonical characterization method for linear systems, *i.e.* system modes.

The notion of non-linear normal modes has been first introduced in the 60's by Rosenberg [118]. The idea is then to consider a mode as a structural *vibration in unison* of all DOF, that is to say a synchronous periodic oscillation of the structure. A mode, characterized by its shape, frequency and damping, is here rather described by a transient deformation of the structure over a cycle at its frequency. Another definition was given by Shaw and Pierre using the notion of invariant manifold [119, 102]. The latter notion was for example exploited by Touzé *et al.* [120] to study system reduction.

The numerical methods employed aim at estimating the bifurcation points by following the non-linear evolution of the system modes as function of control parameters – *i.e.* follow branches of the system. The finality is then to obtain the maximum oscillation amplitudes a system can reach from an unstable point and a given evolution. These methods heavily rely on the concept of extrapolation-correction of an initial tangent state to another one as function of the variation of control parameters. Several methods exists to perform the so-called *continuation*, shooting methods for example by Seydel [121], and arc-length methods, used for example by Ferreira [122], Von Groll *et al.* [111] or Cochelin *et al.* [123] for HBM implementations.

The shooting methods aim at finding the frequency and shape evolution of a new system state from a previous one. A cost function is set evaluating the periodicity error by performing a transient analysis over the new conditions. The shooting is then the way of performing the correction. The initial guess is performed using arc-length methods, to follow an actual system branch.

Sinou [124] proposed the Complex Non Linear Modal Analysis, where the Hopf bifurcation points are estimated by extrapolating the system instability. The solution being expressed on an unstable mode, the amplitude and frequency are extrapolated towards the divergence direction. A new tangent state is then computed, checking if the system is still unstable. Indeed, in the case of the generation of a limit cycle, once the system reaches the limit, the instability disappears to yield a stable periodic solution. This method is related the pseudo-simulations proposed in section 3.4 to approach limit cycles using complex mode trajectories.

Kerschen *et al.* [125, 126] provide a very detailed description of the concepts of non-linear normal modes and the state-of-the-art in terms of computational implementation. The non-linear normal modes are computed as a continuation of the linear ones, as function of a global energy increase. A combination of shooting techniques using the pseudo arc-length method is employed. A great effort to use numerical methods in the continuation phase yields promising results for application to large structures.

Other non-linear analyses are the so-called *global search* approaches, aiming at assessing exhaustively the behavior of a non-linear system. Such method was employed by Sarrouy *et al.* [127] where the system is recast in a polynomial form obtained using a least squares method. The resolution is performed by homotopy, giving potentially all solutions to the polynom, thus giving all equilibrium states of the system.

The interest of global methods relies in its exhaustivity. The classical methods are non-linear analysis from a given equilibrium state. If the said state is not known (in the case of non unicity, some solutions may not directly be found), the potential instability or transition towards this state cannot be assessed.

Historically the *cell-to-cell mapping* was introduced by Hsu in the 80's [128] and later refined [129]. This method aimed at mapping the space of all possible states for a given system. The cells discretize the state space where each cell is given a state representative to the whole cell states. Using time simulation from these initial states, the image of each cell can be found, so that only the attraction points are eventually spotted. The method is however extremely costly. Some improvements of the method involved cell interpolation by Tongue [130], a cell exclusion method, as for example used by Georg [131], or interval search methods. In any way, the computational cost of such method becomes very high or even computationally overwhelming from a dozen DOF systems. They are therefore not really a perspective for industrial applications at time of writing.

In this context the thesis objective is to perform an analysis of the brake stability, and an analysis of the limit cycles. The limit cycle analysis is performed in the time domain, as presented in chapters 4 and 5. Analytical methods are developed in section 3.4, based on the complex modes.

3.3.4 Choices made for the thesis

The studies to be performed must characterize the brake system stability (section 3.4.1) and potential limit cycle amplitude (section 4.3). The clear objective of providing design exploitable results must be kept in mind.

Parametric stability computation will be performed around static working points. From the nominal design results, the effect of numerous parameters can be studied. The CMT presented in section 2.3 will be used for design application, where component free/free modes are the control parameters. More classical studies with parameters, such as friction or braking pressure, will also be performed and presented in appendix C, since this was not a stake for Bosch at this stage.

The limit cycle characterization will be performed in the time domain. Full transient simulations will thus be computed as presented in section 4.3. Design oriented exploitations are made using limit cycles decomposition techniques and transient modal identifications, developed in section 4.4 and illustrated in section 4.5. Further insight was sought, linearizing the system around instantaneous operation points, to analyze the transient evolution of the stability, which is the object of section 5.1.

In parallel to time simulations, a non-linear analysis of the normal modes will be carried out. The objective is to detect at which amplitude saturation patterns can be seen when analyzing complex mode trajectories. This study, presented in section 3.4.2 uses so-called *pseudo-cycles*, which are post-processed like proper time simulations. Pseudo-transient stability evolution is thus analyzed.

3.4 Design using frequency methods in application

This section aims at presenting the classical studies performed for the brake squeal analysis at the industrial level, with some insight gained during the thesis, presented in section 3.4.1. From this analysis, classical global parametric analysis are classically performed – this was however outside the thesis objectives, yet performed and briefly presented in appendix C.

Section 3.4.2 presents a novel approached non-linear analysis based on the system complex modes trajectories. The idea is to study a supposed non-linear response assuming the system shows a single mode based response. Non-static stability analyses can then be derived in section 3.4.3.

3.4.1 Analyzing complex modes of the tangent static state

The current industrial practice for squeal simulation is the use of frequency domain modal computations. The industrial brake models are assembled, meshed, and computed in statics for several pressure points, using local frames. From these pressure points, a tangent state is computed following the definitions of section 3.3.1, real and complex modes are then computed. The real modes are computed using the Lanczos algorithm, described in appendix B. The complex modes are computed in the real mode basis, as detailed in appendix A.

The computed modes are called *nominal*, since they correspond to a precise non-reduced computation using all default parameters. From these simulations, the nominal stability can be studied, as described in section 3.3.1. The characterization can concern the shapes, frequencies, energy distribution, *etc.*

The 12 Bar pressure point and disc velocity of $5 rad/s$ is of particular interest to Bosch and is used as a target if not explicitly stated otherwise. System modes are designated by the prefix R for real modes and C for complex mode followed by the number. Component modes are always taken to be real and are thus referred without prefix.

The brake presented in section 1.2 features several unstable modes at its steady state, as plotted in figure 3.16. In particular, 3 modes can be highlighted as significantly unstable, C44, C51 and C84, whose deformations are shown in figures 3.17, 3.18 and 3.19.

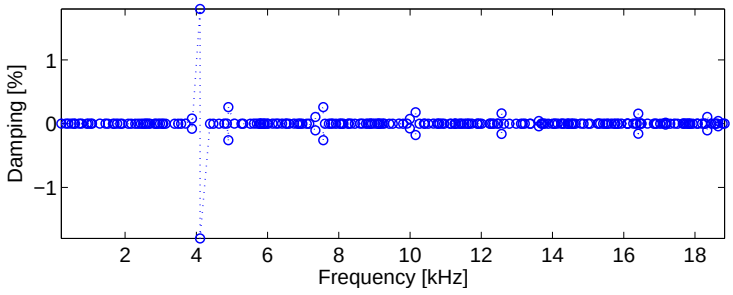
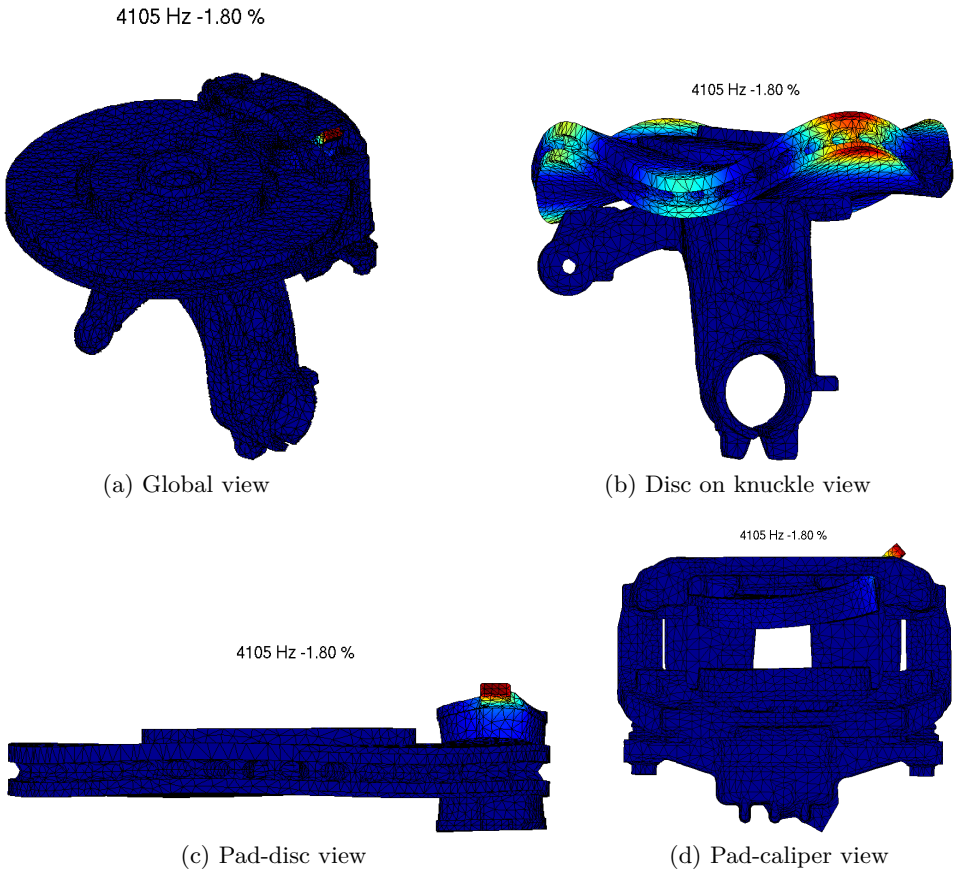


Figure 3.16: Brake stability diagram at 12 Bar

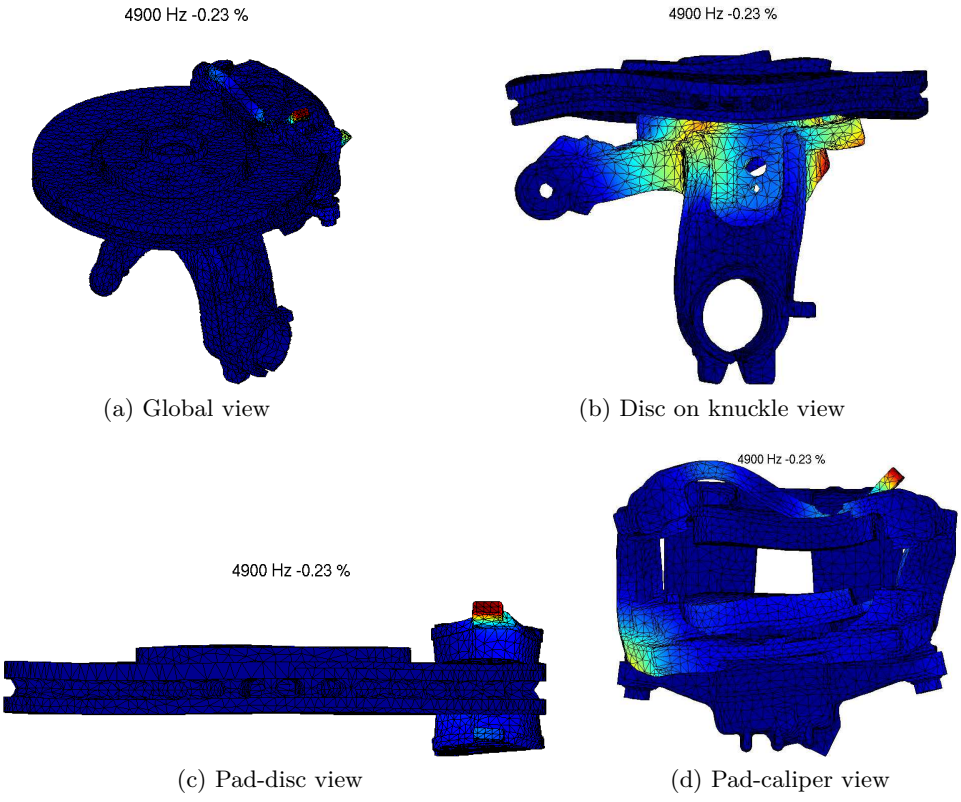
The view of mode C44 in figure 3.17 is decomposed into a global view, and subcomponent views which allows a better understanding of the movements and the existing interactions. Mode C44 features a pad-disc interaction, figure 3.17c, where the outer pad bends at the disc exit, potentially provoking separation. There is low displacements for the caliper and anchor, figure 3.17d. The disc shows a displacement close to a radial mode in figure 3.17b with spacing between nodal diameters close to the pad length.



(a) Global view (b) Disc on knuckle view

(c) Pad-disc view (d) Pad-caliper view

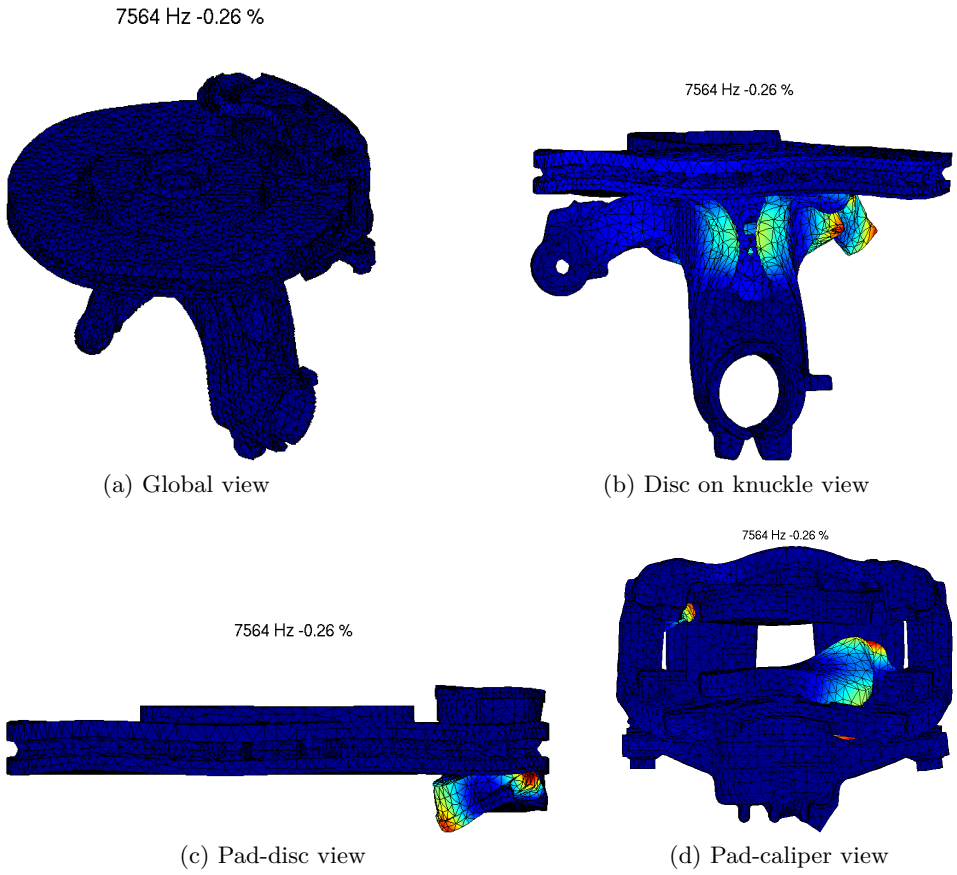
Figure 3.17: Mode C44 shape, colors are re-scaled for each sub-component view



(a) Global view (b) Disc on knuckle view

(c) Pad-disc view (d) Pad-caliper view

Figure 3.18: Mode C51 shape, colors are re-scaled for each sub-component view



(a) Global view

(b) Disc on knuckle view

(c) Pad-disc view

(d) Pad-caliper view

Figure 3.19: Mode C84 shape, colors are re-scaled for each sub-component view

Mode C51 shows pad/disc movements in figure 3.18c similar to mode C44, the global shape is however very different as anchor motion is more significant and thus leads to a stronger pad coupling in figure 3.18d. The disc shows low motion amplitude, figure 3.18b. Significant displacement of the knuckle is certainly related to the anchor coupling.

Mode C84 features an inner pad torsion in figure 3.19d, the anchor and calipers do not show great displacements in figure 3.19d. The disc interaction with the pad is then the cause of the knuckle displacement, in figure 3.19b.

Ranking the strain energy ratios by component already gives more information than just displacements, as shown in figure 3.20. The color code is chosen per component and represents the strain energy ratios from transparent blue to solid dark red.

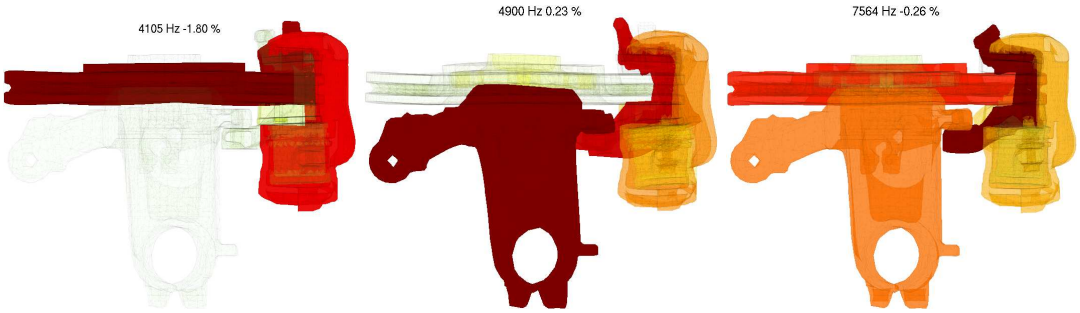


Figure 3.20: Main unstable modes at 12 Bar, C44, C51 and C84, strain energy plotted and ranked by components – from transparent blue (low) to solid dark red (high)

From displacement plots, and figure 3.20 component energy plots, the specificities of each mode can be observed. Mode C44 is the most unstable, it shows large displacements at the pad disc interface, mostly the outer pad rear – where the braking load is lower. The energy is mainly located in three components, the disc, outer pad and caliper.

Mode C51 shows a very different behavior, although the pad shows a large displacement, the energy is located in the knuckle and anchor. This pattern is close to mode C84 where the disc is also solicited.

The strain energy per component plot allows observing coarsely which component can be critical to instability. From a non reduced model, not much can be output from the component wise point of view. This issue is assessed by the CMT, presented in section 2.3 and exploited in chapter 5.

The strain energies plotted by matrices can be interesting. Using the formulation of the CMT method from section 2.3, one can denote for the brake system an elastic matrix, a symmetric stiffness coupling matrix, and an unsymmetric coupling matrix describing the tangent sliding friction. The corresponding plot is given in figure 3.21, where the strain energy ratios are computed for both symmetric matrices.

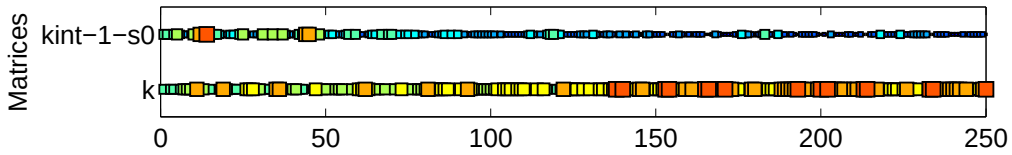


Figure 3.21: System complex mode distribution by matrix, from top to bottom, elastic interaction matrix k_int_s0 (contact and sticking friction) and elastic stiffness k (components). The hotter the color and the larger the marker, the higher the energy fraction

Classical strain energy repartition patterns can be observed as strain energy is present in the interfaces only for lower frequency modes, while higher frequency modes tend to be located only on components. As the frequency increases, modes shapes indeed tend to become localized in specific areas of the system, which is clearly observed here. Mode C44 features some specificity as a lot of its energy is localized at the brake interfaces.

Some insight on the effect of friction coupling on mode C44 can be obtained by looking at its projection to the system real modes, shown in figure 3.22. The projection is trivial as the complex modes are computed in the real mode basis. The modulus of each coefficient of the linear combination of real modes giving the complex mode can then be computed.

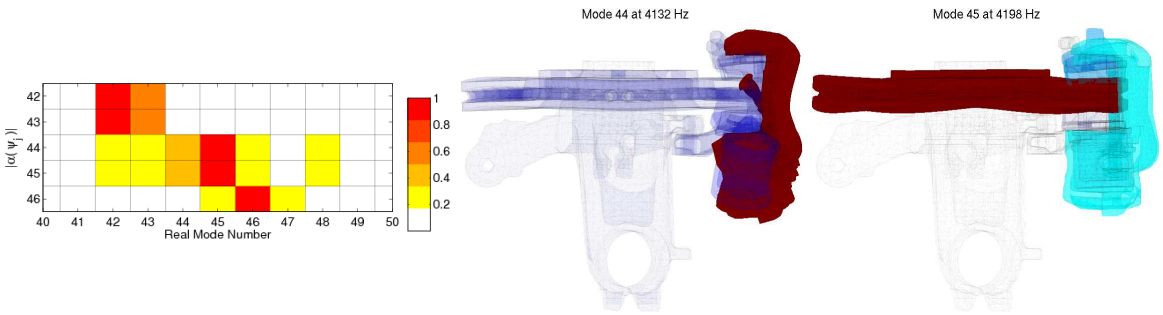


Figure 3.22: Left: Complex mode C44 projection to the system real modes. Right: plot of real modes R44 and R45, main contributors to mode C44 (strain energy ranked and stored by component)

Mode C44 is mainly made of mode R44 showing a high energy in the caliper, and mode R45 showing a high energy in the disc. Friction then couples both modes which strains the outer pad in high energy interaction with both disc and caliper. A lot of these real modes energy is located at the interaction respectively the caliper/pad and the pad/disc. The real modes indeed show that a single component has greater motion than the others in both cases.

This section illustrated typical tools of frequency domain characterization. Multiple complex modes are computed. The Lyapunov theory would state that the relevant ranking is associated with the real part of the eigenvalue. This is however short of relevant design information by providing a nominal only behavior, frequencies and mode shapes associated to a frequency and a damping ratio. Several complex modes are unstable, but only their growth rate is provided. Their ranking in terms of limit cycles amplitude for example is not provided. In other words, there is no *a priori* way of ascertaining what will be the system state in a limit cycle.

3.4.2 Enhanced non-linear analysis of complex modes

The goal of the present non-linear analysis is to have a better understanding of the complex modes computed in section 3.4.1. The role of friction in the expression of the unstable ones needs to be assessed. Their propensity to become prominent in the response, and the attainable amplitudes are the target.

Several unstable modes co-exist, between which a *modal competition* occurs, where the most unstable one should express itself first, depending on their growth rate (or damping ratio). Estimating modal energies and friction works is then an alternative ranking which can contain more information.

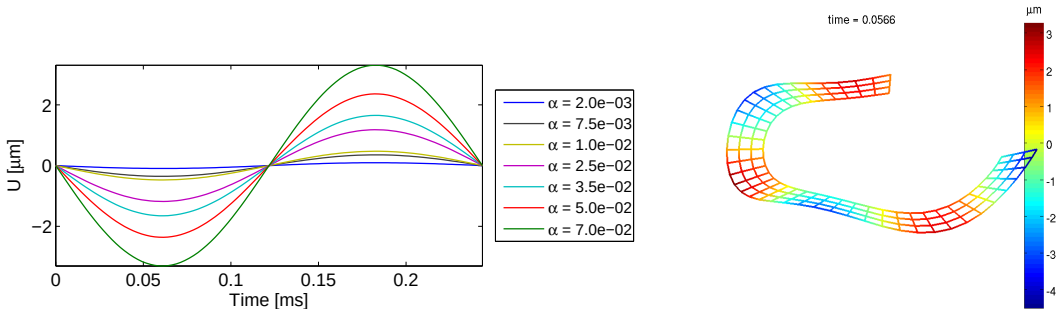
The objective is to quantify the energy levels dissipated in various parts of the system, spotting modes with high friction dissipation levels in the pad/disc contact area. Indeed, they should be the most sensitive ones to the instability.

Damping of a complex mode is normally associated with the natural trajectory given by $\Re\left(\{\psi_j\}\, e^{\lambda_j t}\right)$ which has a decreasing amplitude (damping ratio of the pole) but is not periodic. This notion is clearly not directly related to a dissipation contribution. The proposed idea is thus to consider a cycle of period $T_j = \frac{1}{Im(\lambda_j)} = \frac{1}{\omega_j}$ and trajectory of amplitude α_j

$$\{\Psi_j(t)\} = \Re\left(\alpha_j\, \{\psi_j\}\, e^{i\omega_j t}\right) \tag{3.59}$$

The trajectory defined in equation (3.59) is the base of the generation of *pseudo time simulations*, where it is assumed that only one specific mode is present in the response. Neglecting all other modes, the structure time response would be given for a stabilized self excited vibration by equation (3.59). This should be true at least during the initial divergence from the steady state. The concept can be related to the Complex Non Linear Modal Analysis developed by Sinou [124], where complex modes are taken as the base of a continuation algorithm to compute Hopf bifurcations and limit cycles.

The application is illustrated for mode C44 pseudo-simulations. Figure 3.23 illustrates the pseudo-trajectory based on mode C44, plotted on an observation mesh of the disc in figure 3.23b. The initial coefficient range based on a mass normalized complex mode trajectory is given in figure 3.23a. The range of vibration amplitude is from 0 to 4 μm as function of the coefficient amplitude, which complies with the deformation values that can be experimented with brake systems.



(a) Displacement at a characteristic node

(b) Scaled displacement at maximum amplitude for $\alpha = 7.10^{-2}$

Figure 3.23: Disc displacement levels as function of the modal C44 amplitude α

Clear non-linear effects on the contact-friction forces are observed. The contact forces evolution is of course limited by the full separation of contacting parts while a sinusoidal evolution is observed for contact forces far from separation, as illustrated in figure 3.24.

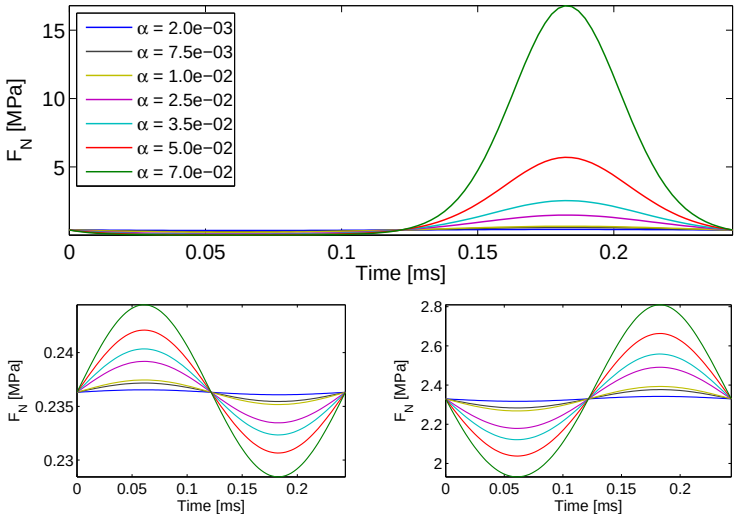


Figure 3.24: Sample contact force oscillation patterns on mode C44 pseudo cycle as function of the vibration amplitude

The obtained braking torque pattern is plotted in figure 3.25, for which a zoom in is provided. The non-linear effects are here illustrated by the brake unloading. Due to separation patterns it hits a threshold as function of the amplitude.

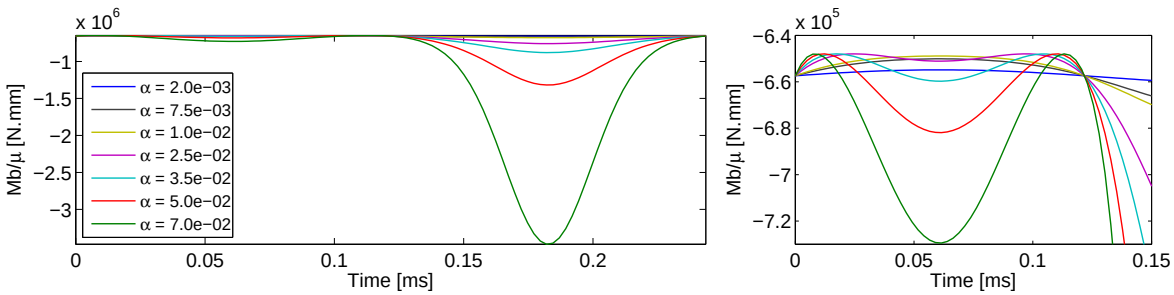


Figure 3.25: Braking torque as function of mode C44 pseudo cycle vibration amplitude, and zoom in on between 0 and 0.15ms

The first idea coming from the evaluation of the pseudo responses is to estimate the work of the friction forces. They are dissipative in general, but are the only source of instability in the system, so that modes prone to respond should feature larger friction works. The pseudo cycles defined by equation (3.59) would correspond to real cycles if the system response was based on a single complex mode.

Noting $\{\psi_j\} = \{\psi_j\}^{(r)} + i\,\{\psi_j\}^{(i)}$, (3.59) yields

$$\{\Psi_j(t)\} = \{\psi_j\}^{(r)} \cos(\omega_j t) - \{\psi_j\}^{(i)} \sin(\omega_j t) \tag{3.60}$$

Considering the friction force work over a pseudo cycle,

$$W(\{\Psi_j\}) = \frac{1}{T_j} \int_0^{T_j} f(\{\Psi_j(t)\}) \cdot \left\{\dot{\Psi}_j(t)\right\} dt \tag{3.61}$$

using the definition of the sliding friction matrix $[K_U(t)]$, for a given tangent contact state $k_{ci}(t)$, the friction force work can be expressed as

$$W(\{\Psi_j\}) = \frac{1}{T_j} \int_0^{T_j} \left([K_U(t))\right] \{\Psi_j(t)\}) \cdot \left\{\dot{\Psi}_j(t)\right\} dt \tag{3.62}$$

which, using (3.60), yields

$$\begin{aligned} W(\{\Psi_j\}) = \quad & -\tfrac{1}{T_j} \textstyle\int_0^{T_j} \left([K_U(t)]\,\{\psi_j\}^{(r)}\right) \cdot \{\psi_j\}^{(r)}\, \omega_j \cos(\omega_j t) \sin(\omega_j t) dt \\ & +\tfrac{1}{T_j} \textstyle\int_0^{T_j} \left([K_U(t)]\,\{\psi_j\}^{(i)}\right) \cdot \{\psi_j\}^{(i)}\, \omega_j \cos(\omega_j t) \sin(\omega_j t) dt \\ & -\tfrac{1}{T_j} \textstyle\int_0^{T_j} \left([K_U(t)]\,\{\psi_j\}^{(r)}\right) \cdot \{\psi_j\}^{(i)}\, \omega_j \cos^2(\omega_j t) dt \\ & +\tfrac{1}{T_j} \textstyle\int_0^{T_j} \left([K_U(t)]\,\{\psi_j\}^{(i)}\right) \cdot \{\psi_j\}^{(r)}\, \omega_j \sin^2(\omega_j t) dt \end{aligned} \tag{3.63}$$

The time dependence of $[K_U]$ makes this integral difficult to compute. Indeed, the friction force work depends on the level of contact, so that the expression could be cast as depending on the positive part of friction. In a first approach however, and estimated as valid for small cycle amplitudes, the time dependence of $[K_U]$ is neglected. The approximation made thus considers the nominal contact state to be constant, and corresponds to the sliding perturbations formulation presented in section 3.2.5. This allows a direct estimation of the work of friction for a given complex mode,

$$W(\{\Psi_j\}) \simeq \frac{\omega_j^2}{2} \left[\left([K_U(q_{stat})]\,\{\psi_j\}^{(i)}\right) \cdot \{\psi_j\}^{(r)} - \left([K_U(q_{stat})]\,\{\psi_j\}^{(r)}\right) \cdot \{\psi_j\}^{(i)}\right] \tag{3.64}$$

A loss factor,

$$\eta\left(\{\Psi_j\}\right) = \frac{2W(\{\Psi_j\})}{\Re(\{\Psi_j\}^T)\,[K]\,\Re(\{\Psi_j\})} \tag{3.65}$$

can be computed as an indicator of the role of friction for each mode, assuming the contact states constant over a cycle.

Equation (3.65) is very cheap computationally, and can give *a priori* estimations for the evaluation of the friction work for each complex mode pseudo cycle.

In application to the industrial brake, equation (3.65) yields the results presented in figure 3.26, where the friction work is estimated using results in absolute values.

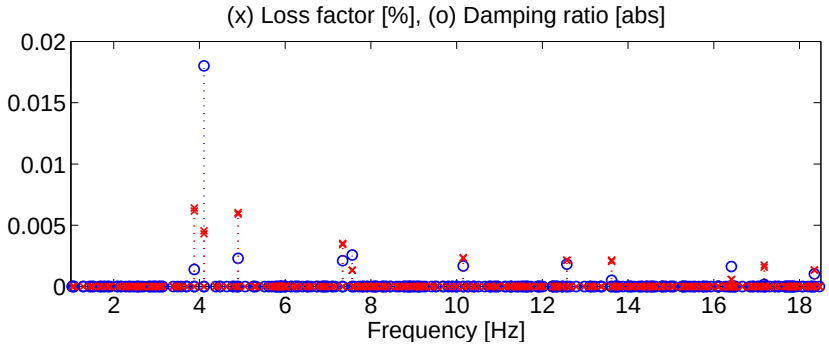


Figure 3.26: Loss factor due to friction in percent against the modal damping ratios in absolute values

It can be observed that only unstable modes generate potentially significant friction works over their cycle. The stability ratios are not in full coherence with the stability plot from figure 3.16 in terms of ranking. Indeed mode C44 which has the highest damping ratio only shows the third loss factor estimation. The constant contact state assumption seems then not to be valid.

To produce better results, the evolution of the contact stiffness must be taken into account. This is easily performed numerically, by actually producing pseudo time simulations. For a given time step, the system states are expressed over a sinusoidal cycle at the complex mode frequency and for a parametered amplitude α_j applied to the shape of the complex mode trajectory given in equation (3.59).

Any time simulation post treatment is then possible, as presented in chapters 4 and 5. In particular the evolution of friction forces as function of the pseudo-time evolution can be computed for every unstable mode cycles, using a basic numerical integration based on equation (3.61). A loss factor is computed,

$$\eta\left(\{\Psi_j\}, \alpha_j\right) = \frac{W\left(\alpha_j \{\Psi_j\}\right)}{\bar{E}_p} \tag{3.66}$$

where $\bar{E}_p$ is the mean strain energy computed over a cycle.

The application to the industrial brake is presented in figure 3.27. It can be noticed during the computation that the friction force work is positive over the cycles. The loss factor in this case is then negative (*i.e.* a gain). All tested modes are then prone to generate instability as the friction will increase the system mechanical energy over time.

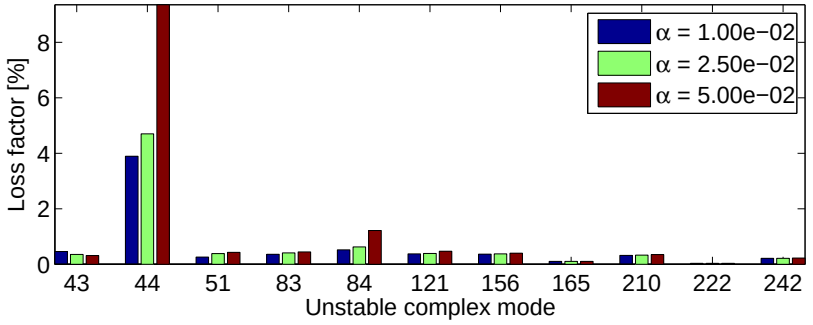


Figure 3.27: Absolute loss factors computed over a pseudo-time simulation based on the unstable complex modes trajectory, for normalized amplitude factors

The results based on pseudo-time integration are still computationally cheap to compute, and will provide more precise results. It can be seen in figure 3.27 that taking the contact pressure variations yields much different results. The effect of amplitude in the mode ranking is important, highlighting the non-linear behavior of the brake.

Some unstable modes show a loss factor decrease as function of the amplitude, although some other only show an increase, this can be linked to separation patterns occurring more or less severely for each mode as function of the amplitude.

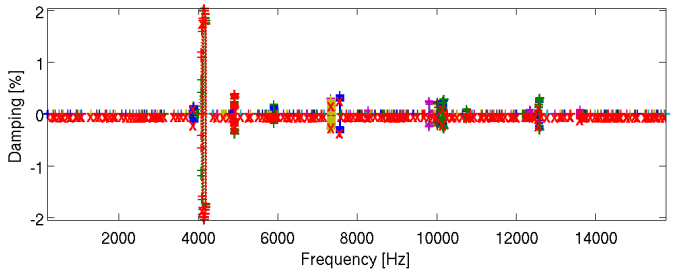
It can eventually be observed that mode C44 is the mode featuring the most work from friction, in correlation with the stability results of figure 3.16, its characterization as the most unstable system mode is then verified by two separate means.

3.4.3 Stability analysis for non static states

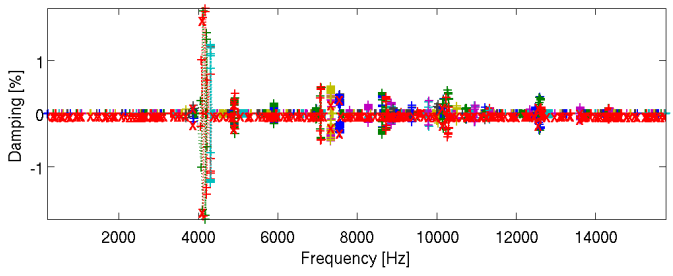
An enrichment of the friction force work estimation is the full re-analysis of the system over the pseudo-trajectory. Since it has been proven that the contact variation cannot be neglected to evaluate the friction work, the idea is here to evaluate the variation of the system tangent state – then of its eigenvalues and its stability – over the pseudo-simulations and the consequent contact stiffness variation occurring. This will be compared to instant tangent state evolutions that can be obtained from actual time simulations in chapter 5.

For all these trajectories, the pseudo instantaneous tangent states can be evaluated and new complex modes can be computed. The evolution of the stabilities and instabilities can therefore be assessed. The following results focus on two sample amplitudes, $\alpha = 10^{-2}$ and $\alpha = 3.5.10^{-2}$. This gives a maximum displacement amplitude of respectively $0.7\mu m$ and $2.3\mu m$.

The relevant amplitude range to consider is obviously between values showing significant variations and unrealistic values. Stability diagrams as function of time are plotted for both amplitudes in figure 3.28. First amplitude in figure 3.28a shows a global conservation of complex modes with great damping variations. Mode C44 shows a variation from -2% to 0%. A few modes in the 6-10kHz bandwidth start to become really unstable for higher amplitudes as shown in figure 3.28b.



(a) $\alpha = 10^{-2}$



(b) $\alpha = 3.5.10^{-2}$

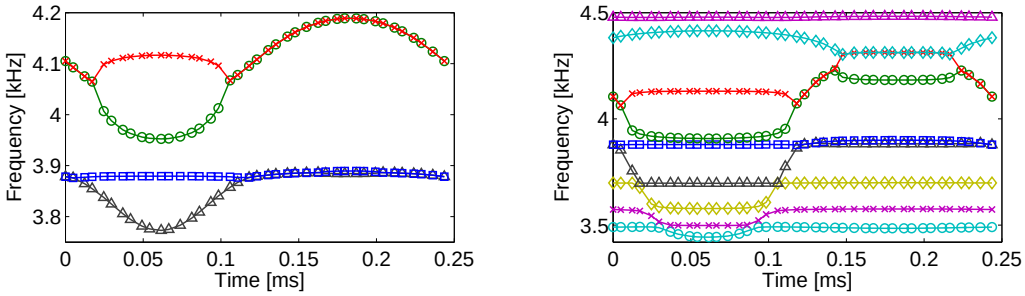
Figure 3.28: Pseudo dynamic stability diagram for two sample amplitudes of mode C44 pseudo cycle

To analyze these results in more details, one will show zoom-ins on the evolution of frequencies as function of the time (position in the pseudo-cycle) and the root locus (position of the poles in the complex plane). Modes C44, C51 and C84 being the most unstable in the static state, they are studied more precisely.

Figure 3.29, 3.30a and 3.32a shows the evolution of mode C44 and its neighbors. Depending on the position in the cycle, one has a pair of poles with nearly equal frequencies and opposite damping or two distinct frequencies and very low damping. The transition between these two states can be seen as coalescence/separation in the frequency/time plot. The system is only unstable in the coalesced state.

Even for low amplitudes in figures 3.29a and 3.30a, a good part of the cycle is associated with a stable tangent state.

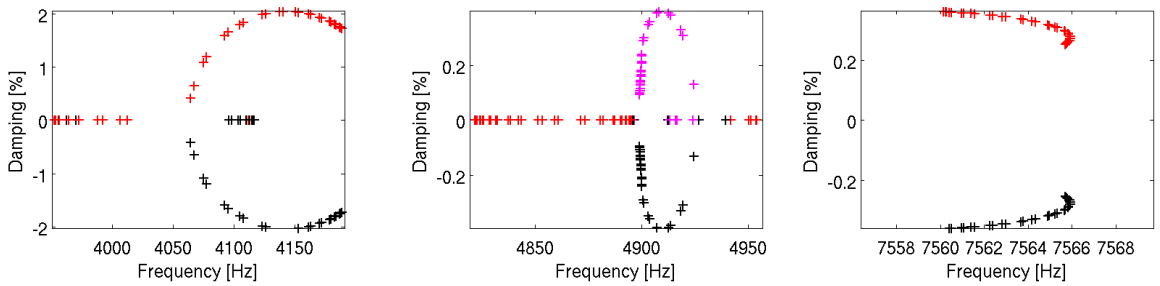
Coalescence is clearly seen even for low amplitude which leads to a stabilization for example in figure 3.30a. Stable/unstable transitions occur for low vibration amplitudes and further coalescence patterns are observed for higher amplitudes, as seen between figures 3.29a and 3.29b.



(a) $\alpha = 10^{-2}$ (b) $\alpha = 3.5.10^{-2}$

Figure 3.29: Frequency evolution of modes around mode C44 (4.2kHz) over an oscillation based on mode C44 pseudo cycle

Over a cycle based on mode C44, the unstable mode pseudo-instant frequency varies, as clearly seen in figure 3.29b. The coalescence pattern of figure 3.29 can be linked to stability transitions in figure 3.30.

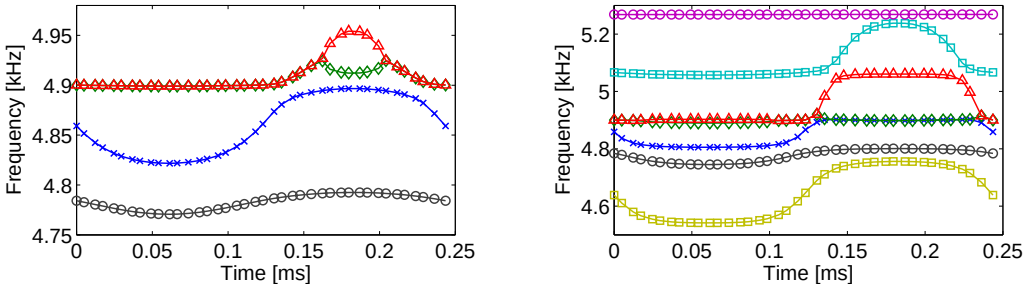


(a) Modes C44 to C45 (b) Modes C50 to C52 (c) Modes C84 to C85

Figure 3.30: Pole evolution of sample modes as function of an oscillation over a pseudo cycle of mode C44 for amplitude $\alpha = 10^{-2}$

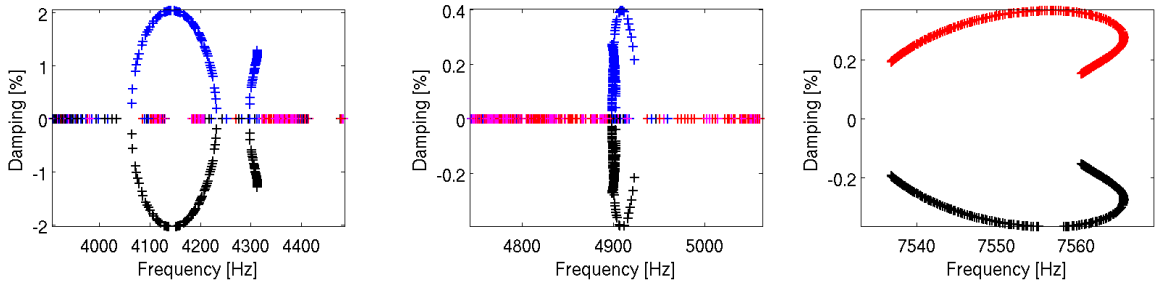
From an initially unstable state at -2%, stabilization occurs over mode C44 cycle. Once coalescence occurs with the lower frequency mode, mode C44 shape transits to a stable state. it becomes unstable again at the end of the cycle.

Mode C51, shows the same type of behavior, to a smaller scale. Figure 3.32b shows that the stability transition happens for $\alpha = 3.5.10^{-2}$ when coalescence and coupling occur in figure 3.31b



(a) $\alpha = 10^{-2}$ (b) $\alpha = 3.5.10^{-2}$

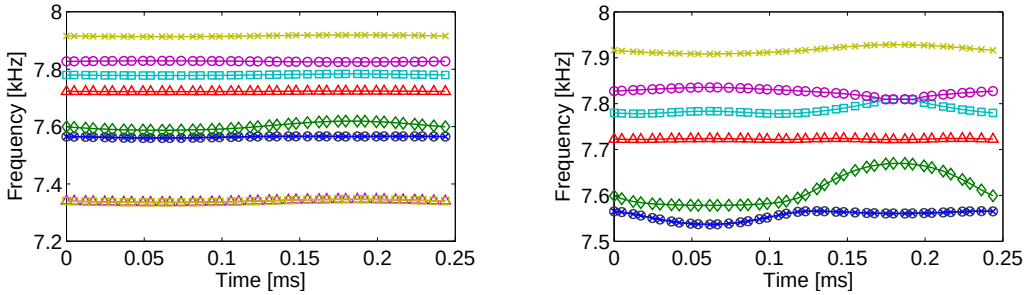
Figure 3.31: Frequency evolution of modes around mode C51 (4.9kHz) over an oscillation based on mode C44 pseudo cycle



(a) Modes C44 to C47 (b) Modes C49 to C52 (c) Modes C84 to C85

Figure 3.32: Pole evolution of sample modes as function of an oscillation over a pseudo cycle of mode C44 for amplitude $\alpha = 3.5.10^{-2}$

Mode C84 stays fully unstable during the whole pseudo-cycle. Figure 3.33 shows that no coalescence patterns happen although some frequency variations can be seen.



(a) $\alpha = 10^{-2}$ (b) $\alpha = 3.5.10^{-2}$

Figure 3.33: Frequency evolution of modes around mode C84 (7.5kHz) over an oscillation based on mode C44 pseudo cycle

The study of mode C44 stability over its own pseudo cycle seems to show that it is not sustainable regarding complete divergence. Such behavior could be a characterization of limit cycle installation. Indeed it was shown that damping would increase if it starts contributing significatively to a time response (through an amplitude increase).

The computation of pseudo-instant modes is more computationally demanding than evaluating the friction force work. The use of very reduced systems such that the Component Mode Models would however allow performing quick and refined re-analyses over several mode cycles in a few hours. This would give precise estimations of the amplitudes of limit cycles under the hypothesis that only one unstable mode responds.

These results are however approached, and need validation through transient simulations. In particular, the contact states used in the analyses presented do not verify mechanical equilibrium. If the amplitude tested becomes too large, the mechanical state can thus become wrong. A perspective of such studies is then to perform mechanical stabilizations at each step of the pseudo-cycles, at least in the vicinity of the contacting areas.

4

Time simulation of structures with non-linearities

Contents

4.1 Time integration of non linear systems

4.1.1 Introduction

Simulation of non-linear systems requires specific handling as standard frequency domain simulations may not provide sufficient information to assess the system behavior. Advanced frequency methods such as the non-linear normal modes defined in section 3.3 require time integration to numerically obtain its cycles. This highlighted the duality between time and frequency methods as a mode can be described by a periodic response over time.

The system to be resolved as function of time t, is formulated here as

$$\begin{cases} [M]\{\ddot{q}(t)\} + [C]\{\dot{q}(t)\} + [K]\{q(t)\} - \{f_c(t)\} - \{f_{NL}(q,\dot{q},t)\} = 0 \\ \{q(0)\}, \{\dot{q}(0)\}, \{\ddot{q}(0)\} \\ \text{B.C. on } \{q\}, \{\dot{q}\} \end{cases} \tag{4.1}$$

where M, C, K are respectively the mass, damping and stiffness matrices, assumed constant in time. q, $\dot{q}$, and $\ddot{q}$ are respectively the displacement, velocity and acceleration response of the structure considered. $f_c(t)$ relates the linear external forces applied to the structure. f_{NL} represents all the non linear loads present in the structure, and depends on time and space fields.

Linearizing systems at a given static – or pseudo-periodic – state automatically limits the results relevance to a certain vibration amplitude range, which cannot be assessed directly. Section 3.4.2 opens the way to compute limitations *a priori*, but actual vibration levels and instability expressions can also be obtained by direct time integration.

Analytical methods, as presented in section 3.3.2 allows estimations but are typically too difficult to implement for the large industrial models, that are the thesis objective. Numerical methods provide trends for realistic models, which is enough regarding the design objectives. The uncertainties scattered at every level of the modeling stages make irrelevant the production of a given analytical value, whose parameterization will be difficult.

Numerical time integration schemes provide ways of resolving a time dependent response by discretizing the time dimension. Methods equivalent to space discretization can therefore be used, being the so-called *finite-element* and *finite-differences* methods.

Finite-differences schemes are based on numerical quadratures for a given time discretization. The time step at a given instant t_n is then function of the previous and following time steps. A time integration time step is said *explicit* if the solution at time t_{n+1} only depends on the previous time steps: $f(t_{n+1}) = \mathcal{G}(t_0, \cdots, t_n)$. An *implicit* time step expresses the solution at time step $n+1$ as function of further time steps as well $f(t_{n+1}) = \mathcal{G}(t_0, \cdots, t_n, \cdots, t_N \geq t_{n+1})$.

Although finite difference schemes have been traditionally used for time integration methods in structural dynamics, more complex schemes can be imagined. In particular the finite element method, well known for spatial discretization may also be applied in the time domain. This opens the way to new discretization methods as the richness of the integration method through time can be controlled, and possibly improved from the finite differences schemes. The optimization between the discretization scale and the formulation order can then be considered for the time domain as well.

Application of time integration methods for structural dynamics shows an interesting history, as their development has been more limited in the 80's and 90's, than in the 70's and 2000's. Once computational power was sufficient to perform numerical integration for mechanical systems in the 60's, a wide range of methods arose, in particular the method developed by Newmark [132], and a large family of methods, from Houblot [133, 134], Wilson [135], or Hilber [136] were developed.

These methods were used for the same applications than in the frequency domain at a time when the question relative to the best method to use was not fully answered. Craig [39] details for example reduction methods for both time and frequency simulations.

Nowadays, the interest of applied research focuses back on time integration mainly for non-linear systems which cannot be solved in the frequency domain. New developments exist, focusing on energy conservation and high order methods able to handle strong linearities, such as the θ-method [87] or the Krenk's scheme [137]. Of course older schemes have also been extensively studied regarding the new applications, for example by Kehnous [138].

4.1.2 First order schemes for state space models

Typical second order systems like equation (4.1) can be reduced to a first order problem, by the mean of a state space formulation, whose canonical form is

$$\dot{q}(t) = [A]\, q(t) + [B]\, \{u(t)\} \tag{4.2}$$

Appendix A shows how a second order equation obtained from classical mechanical systems like equation (4.1) can be transformed to such system.

The *Euler scheme* is a basic finite-differences scheme allowing solving equation (4.2). Equation (4.2) is solved sequentially over a given time discretization spanning the simulation duration T, in time steps $t_0 < \cdots \leq t_n \leq t_{n+1} \leq \cdots < T$, of equal length $h = t_{n+1} - t_n$. The exact response $q_{n+1} = q(t_{n+1})$ is given by

$$q_{n+1} = q_n + \int_{t_n}^{t_{n+1}} \dot{q}(\tau) d\tau \tag{4.3}$$

The numerical method thus aims at approximating the integral in the best way, which is classically obtained by *numerical quadrature*. The velocity on the time interval is then considered as a linear combination of the velocities at time steps t_n and t_{n+1}, introducing a parameter θ,

$$q_{n+1} = q_n + h(1-\theta)\dot{q}_n + h\theta\dot{q}_{n+1} + r_n \tag{4.4}$$

The residue term r_n is ignored in practice. Its computation is however necessary to assess the error made at each time step. For this purpose, one considers the Taylor polynomial expansion of the velocity expressed at times t_n and t_{n+1},

$$\begin{cases} \dot{q}_n = \dot{q}(\tau) + \ddot{q}(t_n - \tau) + q^{(3)}(\tau)\frac{(t_n-\tau)^2}{2} + \mathcal{O}(h^3 q^{(4)}) & \text{(i)} \\ \dot{q}_{n+1} = \dot{q}(\tau) + \ddot{q}(t_{n+1} - \tau) + q^{(3)}(\tau)\frac{(t_{n+1}-\tau)^2}{2} + \mathcal{O}(h^3 q^{(4)}) & \text{(ii)} \end{cases} \tag{4.5}$$

the numerical quadrature then provides the expression of the velocity

$$\dot{q}(\tau) = (1-\theta)\dot{q}_n + \theta\dot{q}_{n+1} + (\tau + \theta h - t_n)\ddot{q}(\tau) + \mathcal{O}(h^2 q^{(3)}) \tag{4.6}$$

whose integration using (4.3) gives the expression of the residue

$$r_n = \left(\theta - \frac{1}{2}\right) h^2 \ddot{q}(\tilde{\tau}) + \mathcal{O}(h^3 q^{(3)}) \tag{4.7}$$

where $\tilde{\tau} \in [t_n \ ; \ t_{n+1}]$.

Using $\theta = \frac{1}{2}$ seems then relevant as the numerical quadrature error expressed in equation (4.7) gains an order. In this case, equation (4.4) is solved by assuming a constant velocity during the time step taken as the average of the velocities bounding the time step, as shown in figure 4.1

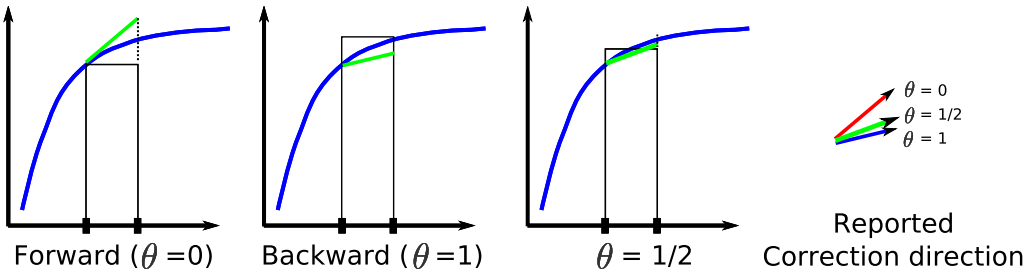


Figure 4.1: Illustration of the interpolation strategy as function of θ for the Euler method

Using $\theta = 0$ makes the scheme explicit and is known as the *Euler forward* method, it is unstable for undamped systems a shown in [139]. Using $\theta = 1$ simplifies the scheme and is known as the *Euler backward* method, it is stable but shows a large numerical damping.

The resolution at each time step requires for implicit cases to solve equation (4.2), as the unknown displacement at time step t_{n+1} is function of its velocity $\dot{q}_{n+1}$ at the same time step. A natural velocity formulation is obtained by combining equations (4.2) and (4.4) such that the system to solve at time step $n+1$ is

$$\begin{cases} \left(h\theta\left[\mathcal{I}\right] + \left[A\right]\right)\dot{q}_{n+1} = \left[A\right]\left(q_n + h(1-\theta)\dot{q}_n\right) + \left[B\right]u_n \\ \\ q_{n+1} = q_n + h(1-\theta)\dot{q}_n + h\theta\dot{q}_{n+1} \end{cases} \tag{4.8}$$

allowing to solve $\dot{q}_{n+1}$, then q_{n+1}.

Recent developments [137, 140, 141, 142, 143] have been introduced to improve time integration of state space systems, by the introduction of finite-element schemes. The philosophy is to directly integrate the state space vector using finite-element shapes in the time domain. Higher order time interpolation functions can thus be used, widening the integration schemes potential. In particular, the state space formalism allows integrating separately displacement and velocity, thus relaxing the mathematical derivative relationship between the two variables. A better precision and stability is commonly obtained for such schemes.

Krenk [137] provides a promising scheme featuring energy control. The main idea is to integrate by parts and separately the displacement and velocity of the expanded state-space vector. From equation (4.2), the state space equation writes, using $v = \dot{q}$,

$$\begin{bmatrix} C & M \\ M & 0 \end{bmatrix}\begin{Bmatrix} \dot{q} \\ \dot{v} \end{Bmatrix} + \begin{bmatrix} K & 0 \\ 0 & -M \end{bmatrix}\begin{Bmatrix} q \\ v \end{Bmatrix} = \begin{Bmatrix} f_{ext}(t) \\ 0 \end{Bmatrix} \tag{4.9}$$

From equation (4.9), displacement q and velocity v are considered separately to integrate the state-space vector, hence the distinction between $\dot{q}$ and v that are now two distinct variables. Integrating equation (4.9) can then be performed only considering approximations on the state-space variables q and v themselves.

The integration of $\dot{q}$ and $\dot{v}$ is performed exactly, writing

$$\begin{cases} \int_{t_n}^{t_{n+1}} \dot{q}(\tau)d\tau = \Delta q = q_{n+1} - q_n \\ \\ \int_{t_n}^{t_{n+1}} \dot{v}(\tau)d\tau = \Delta v = v_{n+1} - v_n \end{cases} \tag{4.10}$$

Equation (4.10) requires the resolution of q_{n+1} and v_{n+1}, whose integration is performed by parts, choosing to derive the whole function and integrating 1 as $\int 1 d\tau = (t - t_n - \frac{h}{2})$,

$$\begin{cases} \int_{t_n}^{t_{n+1}} 1.q(\tau)d\tau = \left[(t - t_n - \frac{h}{2})q(\tau)\right]_{t_n}^{t_{n+1}} - \int_{t_n}^{t_{n+1}}(t - t_n - \frac{h}{2})\dot{q}(\tau)d\tau \\ \\ \int_{t_n}^{t_{n+1}} 1.v(\tau)d\tau = \left[(t - t_n - \frac{h}{2})v(\tau)\right]_{t_n}^{t_{n+1}} - \int_{t_n}^{t_{n+1}}(t - t_n - \frac{h}{2})\dot{v}(\tau)d\tau \end{cases} \tag{4.11}$$

Using equation (4.9) to substitute the derivates appearing in the integrals of equation (4.11) writes

$$\begin{cases} \int_{t_n}^{t_{n+1}} q(\tau)d\tau = \left[(t-t_n-\frac{h}{2})q(\tau)\right]_{t_n}^{t_{n+1}} - \int_{t_n}^{t_{n+1}}(t-t_n-\frac{h}{2})v(\tau)d\tau \\ \\ \int_{t_n}^{t_{n+1}} v(\tau)d\tau = \left[(t-t_n-\frac{h}{2})v(\tau)\right]_{t_n}^{t_{n+1}} - M^{-1}\int_{t_n}^{t_{n+1}}(t-t_n-\frac{h}{2})(Kq(\tau)+Cv(\tau)-f_c)d\tau \end{cases} \tag{4.12}$$

thus introducing the mechanical equations in the scheme formulation.

A finite element interpolation is then performed for q and v. Assuming a linear variation of q and v in the time interval ($\mathcal{P}1$ shape functions),

$$\begin{cases} q(\tau) \simeq \frac{t_{n+1}-\tau}{h}q_n + \frac{\tau-t_n}{h}q_{n+1} \\ \\ v(\tau) \simeq \frac{t_{n+1}-\tau}{h}v_n + \frac{\tau-t_n}{h}v_{n+1} \end{cases} \tag{4.13}$$

The integrals of equation (4.12) can now be computed using (4.13), yielding

$$\begin{cases} \int_{t_n}^{t_{n+1}} q(\tau)d\tau \simeq \frac{h}{2}(q_{n+1}+q_n) - \frac{h^2}{12}\Delta v \\ \\ \int_{t_n}^{t_{n+1}} v(\tau)d\tau \simeq \frac{h}{2}(v_{n+1}-v_n) - \frac{h^2}{12}M^{-1}(K\Delta q + C\Delta v - \Delta f_c) \end{cases} \tag{4.14}$$

Introducing equations (4.10) and (4.14) in the state-space equation (4.9) eventually yields the Krenk integration scheme

$$\begin{bmatrix} C+\frac{h}{2}K & M-\frac{h^2}{12}K \\ M-\frac{h^2}{12}K & -\frac{h}{2}M-\frac{h^2}{12}C \end{bmatrix} \begin{Bmatrix} q_{n+1} \\ v_{n+1} \end{Bmatrix} = \begin{bmatrix} C-\frac{h}{2}K & M-\frac{h^2}{12}K \\ M-\frac{h^2}{12}K & \frac{h}{2}M-\frac{h^2}{12}C \end{bmatrix} \begin{Bmatrix} q_n \\ v_n \end{Bmatrix} + \begin{Bmatrix} \int_{t_n}^{t_{n+1}} f_c d\tau \\ -\frac{h^2}{12}\Delta f_c \end{Bmatrix} \tag{4.15}$$

The scheme summarized in equation (4.15) is deemed promising in respects to its very good mathematical properties in terms of precision and stability, using evaluation methods presented in section 4.1.5 for the Newmark scheme. Its good properties mainly come from the fact that displacement and velocities are computed separately, allowing a fourth order accuracy for linear models. Krenk's formalism also allows the introduction of so-called algorithmic (or numerical – in section 4.1.5) damping in an optimized and controlled way.

In a more general context, finite element schemes in the time domain allows relaxing continuity properties in new ways. The Discontinuous Galerkin (DG) scheme, introduced by Hulber *et al.* [144] allows in particular discontinuities between time steps. The application is of interest as the system representation can be changed between two time steps, provided a projection exists between both models. The target application is then adaptative meshing used for example by D. Aubry *et al.* [140, 141, 142] for simulations of wave propagation in elastic 3D solids. The compromise between model size and mesh refinement is obtained by finely meshing only local areas at propagation fronts. The mesh potentially changes between each time step, which generates displacement discontinuities, well handled by DG schemes.

To handle discontinuities between two time steps, the classical formulation of the mechanical problem must be completed. Coming back to a continuous formulation for an undamped system,

$$\left\{\begin{array}{l} \displaystyle\int_{\Omega\times T_n} \rho\ddot{q}(x,t)w_1(x,t)dSdt + \int_{\Omega\times T_n} tr(\sigma(q(x,t)\epsilon(w_1(x,t)))dSdt + \int_{\Omega} \rho[q]_n w_1(t_n)dS \\ \qquad\qquad = \displaystyle\int_{\Omega\times T_n} f_c(x,t)w_1(x,t)dSdt \\ \displaystyle\int_{\Omega\times T_n} tr(\sigma(\dot{q}(x,t)-v(x,t))\epsilon(w_2(x,t)))dSdt + \int_{\Omega\times T_n} tr(\sigma([\dot{q}]_n))\epsilon(w_2(t_n)))dSdt = 0 \end{array}\right. \tag{4.16}$$

with w_1 and w_2 are the test functions of the weak formulation.

It can be observed the need to take into account the acceleration jump to balance inertial forces between two time steps, and the need to append an equation handling the velocity jumps.

A more general formalism than for the Krenk scheme can be setup; finite element discretizations are defined for both space and time dimensions. The most simple shape functions will however separate the space and time dimensions, so that the displacement u of a point located at x at instant t will be interpolated using space functions N_i and time shape functions V_j,

$$q(x,t) \simeq \sum_{i,j} N_i(x)V_j(t) \tag{4.17}$$

This allows separating space and time integrations of the set of equations (4.16). Equation (4.16) can then be integrated classically in the space domain. Assuming u_x the space discretized displacement on the volume, and using $\mathcal{P}1$ shape functions for the sake of simplicity in the method illustration provided in this section, one assumes the time interpolation as

$$q_x(t) = \frac{t_{n+1}-t}{h}q_{x,n} + \frac{t-t_n}{h}q_{x,n+1} \tag{4.18}$$

Due to the discontinuity handling, velocity and displacement are treated separately, so that both of them are interpolated using the shape functions given in equation (4.18). Their derivative in time is of course obtained by the derivation of equation (4.18). After all integrations are performed, both equations are taken for test functions being null first at time t_n, then at time t_{n+1}, so that 4 equations can be exploited. Indeed, displacement and velocity at both times t_n and t_{n+1} must be solved because of the discontinuities,

$$\left[\begin{array}{cccc} \frac{h}{6}K & \frac{h}{3}K & -\frac{1}{2}M & \frac{1}{2}M \\ \frac{h}{3}K & \frac{h}{6}K & \frac{1}{2}M & \frac{1}{2}M \\ -\frac{1}{2}K & \frac{1}{2}K & -\frac{h}{6}K & -\frac{h}{3}K \\ \frac{1}{2}K & \frac{1}{2}K & -\frac{h}{3}K & -\frac{h}{6}K \end{array}\right] \left\{\begin{array}{c} q_n \\ q_{n+1} \\ v_n \\ v_{n+1} \end{array}\right\} = \left\{\begin{array}{c} \frac{h}{3}f_{n+1} + \frac{h}{6}f_n \\ Mv_n^- + \frac{h}{3}f_n + \frac{h}{6}f_{n+1} \\ 0 \\ Kq_n^- \end{array}\right\} \tag{4.19}$$

A first inconvenient is clearly seen here as the system is four times larger than the initial one. A bold application of such scheme seems then counter productive for industrial applications, where size matters. The raw formulation presented in equation (4.19) can however be tweaked to lighten the equations. Equation (4.19) is indeed equivalent to

$$\left[\begin{array}{cccc} K & 0 & -\frac{h}{6}K & -\frac{h}{6}K \\ 0 & K & -\frac{h}{2}K & -\frac{h}{2}K \\ \frac{h}{6}K & -\frac{h}{6}K & M & 0 \\ \frac{h}{2}K & \frac{h}{2}K & 0 & M \end{array}\right] \left\{\begin{array}{c} q_n \\ q_{n+1} \\ v_n \\ v_{n+1} \end{array}\right\} = \left\{\begin{array}{c} Kq_n^- \\ Kq_n^- \\ f_n - f_{n+1} + Mv_n^- \\ f_n + f_{n+1} + Mv_n^- \end{array}\right\} \tag{4.20}$$

This rearrangement is more favorable not only because null blocks appear for matrix/vector products, but also because this allows the generation of smaller systems based on a predictor/corrector formulation of (4.20). Indeed, the displacement terms in equation (4.20) are only dependent on the velocity,

$$\begin{cases} q_n = q_n^- - \frac{h}{6}(v_{n+1} - v_n) \\ \\ q_{n+1} = q_n^- + \frac{h}{2}(v_{n+1} + v_n) \end{cases} \tag{4.21}$$

Using predictions in displacement, the velocities are solved using the last two equations of (4.20),

$$\begin{bmatrix} M + \frac{h^2}{6}K & \frac{2}{3}M \\ \frac{h^2}{3}K & M + \frac{h^2}{6}K \end{bmatrix} \begin{Bmatrix} v_n \\ v_{n+1} \end{Bmatrix} = \begin{Bmatrix} f_n - f_{n+1} + Mv_n^- - \frac{h}{6}K(q_n - q_{n+1}) \\ f_n + f_{n+1} + Mv_n^- - \frac{h}{6}K(q_n + q_{n+1}) \end{Bmatrix} \tag{4.22}$$

A lot of implementation variants of the prediction-correction scheme from equations (4.21) and (4.22) are possible, some of which are detailed in [140]. An algorithm is given by Chien *et al.* [145], along with performance comparisons with usual finite differences schemes.

The DG schemes are usually formulated for linear elastic undamped systems. The interest was then focused on spatial discontinuities adapted to an adaptative remeshing. The application targeted here is however different, as non linear laws are simulated for a constant space over time. The ability for such scheme to handle strong non linearities as generated by contact-friction Lagrange constraints can then be studied. Such application was initiated by Hulbert [146] for sample impact applications.

More recent developments by Mancuso *et al.* aim at implementing the DG schemes for more general classes of non-linear systems. Two main ideas are developed in [143] giving the philosophy of the DG application to non-linear systems. First the prediction-correction schemes can be adapted from the standard formulations to use Newton-Raphson techniques. Second, the integration of non-linear forces terms must be carefully handled. Bonelli [147] extends the latter idea by performing three points Gauss-Legendre quadratures for the integrals of non-linear forces.

The main drawbacks of time integration schemes based on state space systems is mainly an increase of the system size by two due to the first order formulation, which can become a great performance problem when dealing with industrial models. The lack of extended studies to non-linear models for finite element schemes also makes it an unknown for the squeal application targeted. They are therefore not retained. Improved stability and precision properties provided by finite-elements schemes are nevertheless interesting and could become a very relevant choice in the future.

4.1.3 The θ-method, a formalism adapted to non-smooth dynamics

The term of *non-smooth dynamics* is well related to the mathematical difficulties arising from contact-friction formulations based on Lagrange formulations. Indeed the use of constraints alters the continuity properties of the signals to integrate. In particular the acceleration cannot be considered as continuous everywhere which is for example the case when impacts and stick/slip occur in a non-regularized way. The θ-method application range is however more general than contact only. It is the base of the *non-smooth contact dynamics method* developed by Moireau and Jean [87].

The regular formulation from equation (4.1) is replaced by a more general second order mechanical equation, assuming the acceleration as a distribution, so that the evolution of the system between time steps t_n and t_{n+1} is written

$$\int_{t_n}^{t_{n+1}} [M]\, \ddot{q} d\tau = -\int_{t_n}^{t_{n+1}} \left([K]\, q(\tau) + [C]\, \dot{q}(\tau)\right) d\tau + \int_{t_n}^{t_{n+1}} \left(f_c + f_{NL}(q, \dot{q}, t)\right) d\tau \tag{4.23}$$

Since the irregularities can occur within a time step, one only considers the integral $\int_{]t_n,t_{n+1}]} [M]\,\ddot{q}d\tau$ over a time step, writing

$$\int_{t_n}^{t_{n+1}} [M]\,\ddot{q}d\tau = [M]\,(\dot{q}_{n+1} - \dot{q}_n) \tag{4.24}$$

For the same reasons, $\int_{]t_n,t_{n+1}]} f_{NL}(q,\dot{q},t)d\dot{q}$ is computed using a mean impulse value, assuming the non-linear forces are known at time step $n+1$

$$\frac{1}{h}\int_{t_n}^{t_{n+1}} f_{NL}(q(\tau),\dot{q}(\tau),\tau)d\tau = f_{NL}(q_{n+1},\dot{q}_{n+1},t_{n+1}) \tag{4.25}$$

Regular forces – strain, viscous, and external – are computed using a numerical quadrature introducing a parameter θ, such that

$$\int_{t_n}^{t_{n+1}} f(\tau)d\tau = h\theta f_{n+1} + h(1-\theta)f_n \tag{4.26}$$

giving the expression of the remaining integrals of equation (4.23)

$$\begin{aligned}\int_{t_n}^{t_{n+1}} ([K]\,q(\tau) + \quad & [C]\,\dot{q}(\tau) + f_c(\tau))d\tau = \\ & h\theta([K]\,q_{n+1} + [C]\,\dot{q}_{n+1} + f_{cn+1}) + h(1-\theta)([K]\,q_n + [C]\,\dot{q}_n + f_{cn})\end{aligned} \tag{4.27}$$

The displacement is integrated by the same mean yielding here the Euler quadrature,

$$q_{n+1} = q_n + h(1-\theta)\dot{q}_n + h\theta\dot{q}_{n+1} \tag{4.28}$$

Practically, assumptions of equation (4.25) requires a prediction and correction strategy. The combination of equations (4.23), (4.25), (4.24), (4.27) and (4.28) yields the practical θ-method scheme,

$$\begin{cases} (M + h\theta C + h^2\theta^2 K)\,\dot{q}_{n+1} = & (M - h^2\theta(1-\theta)K - h(1-\theta)C)\,\dot{q}_n \\ \qquad -hKq_n + h\theta f_{cn+1} + & h(1-\theta)f_{cn} + hf_{NLn+1}(q_{n+1},\dot{q}_{n+1},t_{n+1}) \\ \\ q_{n+1} = q_n + h(1-\theta)\dot{q}_n + h\theta\dot{q}_{n+1} \end{cases} \tag{4.29}$$

whose algorithmic implementation is given in figure 4.2, where the non-linear forces are directly expressed by their Lagrangian formulation, (see section 3.2.1).

The Jacobian used for the correction interactions $[J]$ is here written

$$[J] = [M] + \theta h\,[C] + \theta^2 h^2\,[K] \tag{4.30}$$

Such scheme has been used by Moirot [5] to Lorang [6] to perform squeal simulations on train brakes, or by Raous [148] for friction instability studies. A modification of the contact formulation was performed by Moireau and Jean [87] to allow inelastic impacts. Indeed, the quadrature formulation from equation (4.4) imposes the gap observation (using section 3.2.4 notations)

$$g_{n+1} = C_{NOR}q_{n+1} = C_{NOR}(q_n + h(1-\theta)\dot{q}_n + h\theta\dot{q}_{n+1}) \tag{4.31}$$

which relates gap displacements and velocities at time step $n+1$. This leads to the definition of a gap velocity,

$$\dot{g}_{n+1} = C_{NOR}\dot{q}_{n+1} = \frac{1}{h\theta}(g_{n+1} - g_n) + \frac{\theta-1}{\theta}\dot{g}_n \tag{4.32}$$

If the contact closes at time step n, the gap remains constant for time steps n and $n+1$, while the gap velocity at time step n is not null. This is a problem if $\theta \neq 1$ since the gap velocity at time step $n+1$ cannot be null by equation (4.32). Inelastic contact would then be impossible.

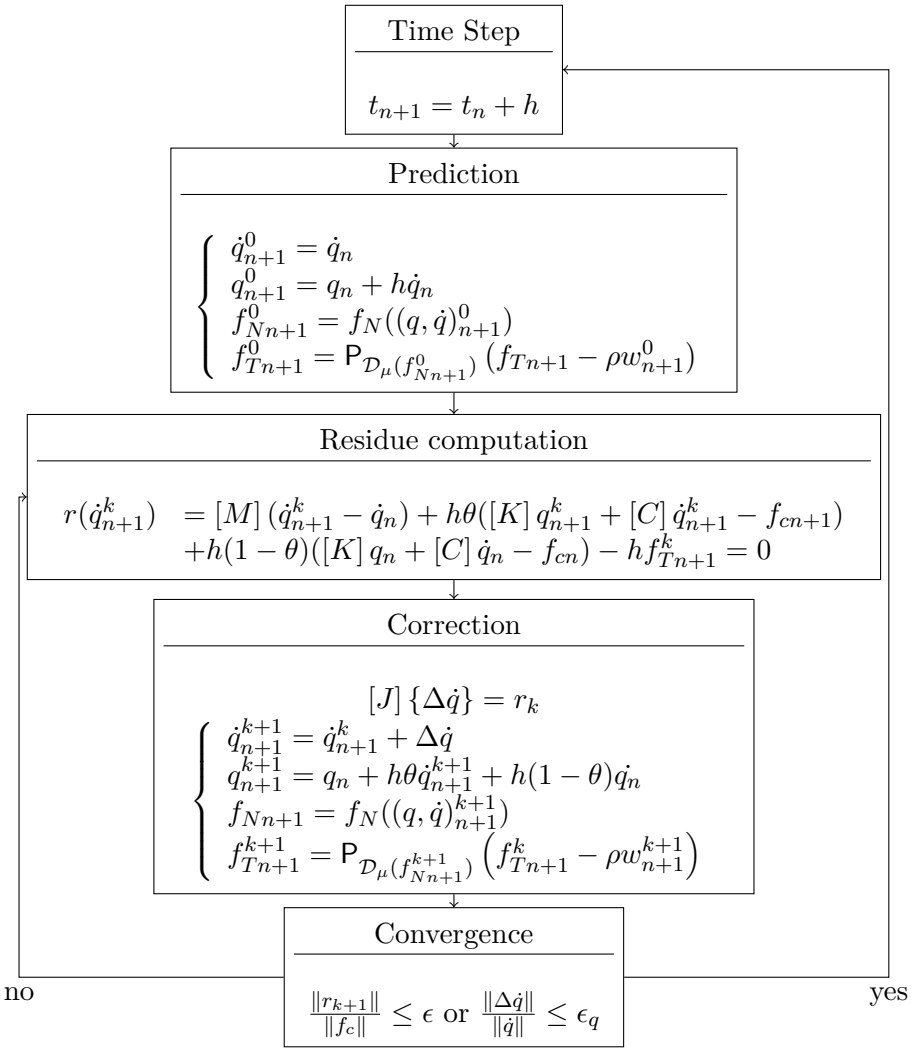


Figure 4.2: Implementation of the θ-method for the Lagrange friction force formulation

The proposed solution is to compute gaps with an offset of time. The gap computed at time step n is then taken at another instant between t_n and t_{n+1}. The new gap $\tilde{g}_{n+1}$ is computed at time $t_{n+1} + h(1-\theta)$, assuming the velocity constant in the interval,

$$\tilde{g}_{n+1} = g_{n+1} + h(1-\theta)\dot{g}_{n+1} \tag{4.33}$$

which from the quadrature (4.4) yields

$$\tilde{g}_{n+1} = \tilde{g}_n + h\dot{g}_{n+1} \tag{4.34}$$

thus the new gap velocity at time step $n+1$ is given by

$$\dot{g}_{n+1} = \frac{1}{h}(\tilde{g}_{n+1} - \tilde{g}_n) \tag{4.35}$$

which does not depend on the velocity at time step n.

Using such gap allows to tackle the problem from equation (4.32), as it is now possible to have a null gap velocity at time step $n+1$ in the conditions of an impact.

The θ-method is well adapted for a Lagrange formulation of contact problems. Numerical damping occurs due to the numerical quadrature unless $\theta = \frac{1}{2}$, as presented by Vola [4].

The approximations not conceded at the level of contact law are conceded at the non-linear forces integration and regularization in fact occurs at the gap computation level. The proposed regularization cannot be assessed physically which can raise potential interpretation issues.

4.1.4 Second order finite differences schemes: the Newmark family

When the system features smooth laws, the acceleration can be defined at each instant, which allows the use of second order time schemes. These schemes can be displacement or acceleration based and are the most common in the literature. The interest of a second order scheme is that classical second order equations like (4.1) can be solved directly.

In the case where contact-friction forces are regularized, using laws presented in section 3.2 for example, these more traditional schemes can be used to integrate the system.

The numerical quadrature performed in section 4.1.2 on the velocity is here performed on the acceleration. The exact displacement and velocity are formulated as function of the acceleration

$$\begin{cases} \dot{q}_{n+1} = \dot{q}_n + \int_{t_n}^{t_{n+1}} \ddot{q}(\tau) d\tau \\ \\ q_{n+1} = h\dot{q}_n + \int_{t_n}^{t_{n+1}} (t_{n+1} - \tau)\ddot{q}(\tau) d\tau \end{cases} \tag{4.36}$$

Two different integrals are then approximated by numerical quadrature, the quadrature for the velocity (resp. displacement) is performed using a parameter γ (resp. 2β)

$$\begin{cases} \dot{q}_{n+1} = \dot{q}_n + h(1-\gamma)\ddot{q}_n + h\gamma\ddot{q}_{n+1} + r_n \\ \\ q_{n+1} = q_n + h\dot{q}_n + h^2\left(\frac{1}{2} - \beta\right)\ddot{q}_n + h^2\beta\ddot{q}_{n+1} + r_n^{'} \end{cases} \tag{4.37}$$

Quadrature errors r_n and $r_n^{'}$ are obtained from the expressions of the acceleration for $\tau \in [t_n \ ; \ t_{n+1}]$ using Taylor polynomial expansions

$$\begin{cases} \ddot{q}_n = \ddot{q}(\tau) + q^{(3)}(\tau)(t_n - \tau) + \mathcal{O}(h^2 q^{(4)}) \\ \\ \ddot{q}_{n+1} = \ddot{q}(\tau) + q^{(3)}(\tau)(t_{n+1} - \tau) + \mathcal{O}(h^2 q^{(4)}) \end{cases} \tag{4.38}$$

The quadrature errors are thus deduced from equation (4.38) as

$$\begin{cases} r_n = \left(\gamma - \frac{1}{2}\right) h^2 q^{(3)}(\tilde{\tau}) + \mathcal{O}(h^3 q^{(4)}) \\ \\ r_n^{'} = \left(\beta - \frac{1}{6}\right) h^3 q^{(3)}(\tilde{\tau}) + \mathcal{O}(h^4 q^{(4)}) \end{cases} \tag{4.39}$$

$\beta = \frac{1}{6}$ and $\gamma = \frac{1}{2}$ is the relevant choice to minimize the error. Section 4.1.5 will however show that this set of parameters is unstable, so that the most commonly used values are $\beta = \frac{1}{4}$ and $\gamma = \frac{1}{2}$.

Looking at equations (4.7) and (4.39), it can be observed that the quadrature errors obtained for the displacement is an order higher for the second order scheme. Provided that the model is *regular enough*, a more precise integration than with Euler schemes is thus obtained.

The resolution at each time step of the mechanical equilibrium from equation (4.1) is commonly expressed as acceleration based, the quadrature formulations from equation (4.39) gives the system to solve at time step $n+1$, provided time step n

$$\begin{cases} \left([M] + h\gamma\,[C] + h^2\beta\,[K]\right)\ddot{q}_{n+1} = & \left(h(1-\gamma)\,[C] + h^2(\frac{1}{2} - \beta)\,[K]\right)\ddot{q}_n \\ & + \left([C] + h\,[K]\right)\dot{q}_n + [K]\,q_n + f_{cn+1} \\ \\ \dot{q}_{n+1} = \dot{q}_n + h(1-\gamma)\ddot{q}_n + h\gamma\ddot{q}_{n+1} \\ \\ q_{n+1} = q_n + h\dot{q}_n + h^2\left(\frac{1}{2} - \beta\right)\ddot{q}_n + h^2\beta\ddot{q}_{n+1} \end{cases} \tag{4.40}$$

allowing to solve $\ddot{q}_{n+1}$ then $\dot{q}_{n+1}$ and q_{n+1}.

The formulation from equation (4.40) is the most simple in linear dynamics. In the case of non-linear dynamics however, displacement based formulations are generally preferred. Indeed, non-linear forces are typically function of the displacement and velocity, as written in equation (4.1).

The non-linear resolution of the mechanical equation from system (4.40) is usually performed by an iterative predictor/corrector scheme. Given the solution at time step n, the prediction is initialized by assuming a null acceleration at time step $n+1$, so that from system (4.40) the predictors q^0_{n+1} and $\dot{q}^0_{n+1}$ are expressed as

$$\begin{cases} q^0_{n+1} = q_n + h\dot{q}_n + h^2(\frac{1}{2} - \beta)\ddot{q}_n \\ \dot{q}^0_{n+1} = \dot{q}_n + h(1-\gamma)\ddot{q}_n \end{cases} \tag{4.41}$$

One considers the displacement correction Δq_{n+1} as the only unknown and can show [139] that velocity and acceleration at time step $n+1$ are given by

$$\begin{cases} \Delta q_{n+1} = q_{n+1} - q^0_{n+1} \\ \dot{q}_{n+1} = \dot{q}^0_{n+1} + \frac{\gamma}{h\beta}\Delta q_{n+1} \\ \ddot{q}_{n+1} = \frac{1}{\beta h^2}\Delta q_{n+1} \end{cases} \tag{4.42}$$

As the system to solve is non linear, an incremental correction will be applied, until the correction at increment N becomes negligible regarding a given tolerance ϵ, so that the displacement correction will rather be expressed as a correction sequence

$$\Delta q_{n+1} = \sum_{k=1}^{N} \Delta q^k_{n+1} \tag{4.43}$$

From iteration k to iteration $k+1$ equation (4.42) can be written in the same way

$$\begin{cases} \Delta q^{k+1}_{n+1} = q^{k+1}_{n+1} - q^k_{n+1} \\ \ddot{q}^{k+1}_{n+1} = \ddot{q}^k_{n+1} + \frac{1}{\beta h^2}\Delta q^{k+1}_{n+1} \\ \dot{q}^k_{n+1} = \dot{q}^k_{n+1} + \frac{\gamma}{h\beta}\Delta q^{k+1}_{n+1} \end{cases} \tag{4.44}$$

The mechanical equation to solve is written as

$$[M]\,\ddot{q}^k_{n+1} + [C]\,\dot{q}^k_{n+1} + [K]\,q^k_{n+1} - f_{cn+1} - f_{NLn+1}(q^k_{n+1}, \dot{q}^k_{n+1}, t_{n+1}) \tag{4.45}$$

A classical interpretation of equation (4.45) is to consider the mechanical resolution at increment $k+1$ as a direct correction of the residue coming from the fact that the solution at increment k does not comply with the mechanical equilibrium. Provided solution q^k_{n+1}, the residue can be defined as

$$r^{k+1}_{n+1} = [M]\,\ddot{q}^k_{n+1} + [C]\,\dot{q}^k_{n+1} + [K]\,q^k_{n+1} - f_{cn+1} - f_{NLn+1}(q^k_{n+1}, \dot{q}^k_{n+1}, t_{n+1}) \tag{4.46}$$

Equation (4.46) is thus iteratively solved by a Newton-Raphson algorithm by computing the derivative of the residue as function of the displacement,

$$[J(q,\dot{q})]_{|\Delta q} = \frac{\partial r(q,\dot{q})}{\partial \Delta q}_{|\Delta q} = \frac{1}{\beta h^2}[M] + \frac{\gamma}{\beta h}[C] + [K] + \frac{\partial f_{NL}(q+\Delta q, \dot{q} + \frac{\gamma}{\beta h}\Delta q)}{\partial \Delta q} \tag{4.47}$$

In this equation the Jacobian is clearly linked to the derivative of the non linear forces at the current state. In the present work, the use of a fixed Jacobian, which corresponds to a modified Newton-Raphson algorithm, is a necessity to avoid repeated factorizations of the Jacobian. A fixed stiffness $[K_\tau]$ associated with the contact forces is thus introduced leading to

$$[J] = \frac{1}{\beta h^2}[M] + \frac{\gamma}{\beta h}[C] + [K] + [K_\tau] \tag{4.48}$$

The use of this fixed Jacobian is a key feature needed for the performance of time simulations shown in this work. The procedure used for its selection will be discussed in section 4.2.2. It should be noted that modifying the Jacobian does not change the solution but only the convergence properties of the scheme.

The system thus formed allows integrating the time response of the initial problem (4.1). The scheme is then referred to the *implicit non linear Newmark scheme* and can be summarized in the process diagram of figure 4.3.

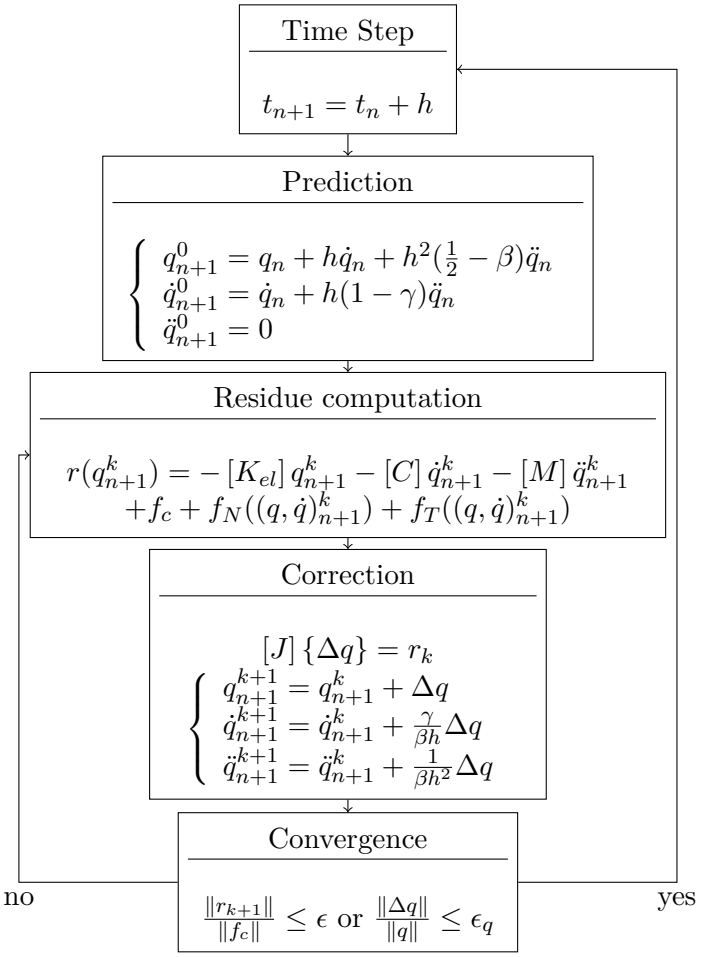


Figure 4.3: Implementation of the modified non linear Newmark scheme

4.1.5 Evaluation of a time integration scheme, application to the Newmark scheme

The schemes presented in sections 4.1.2 and 4.1.4 are the most common schemes used in the literature. Their parameters can however be tuned to obtain specific properties. Besides, their application to non linear systems can generate peculiar behaviors. This needs to be evaluated to control the simulation results. First the mathematical picture of the requirements for an integration scheme to be relevant is given. Second, these notions are exploited in a more mechanical way, which is critical for the results interpretation and the thesis modeling choices.

Consistency and *stability* are two distinct mathematical notions characterizing the relevance of an integration scheme. The consistency is verified if the approximate solution converges towards the continuous solution as the time step is decreased, that is

$$\lim_{h\to 0} \frac{q_{n+1} - q_n}{h} = \dot{q}(t_n) \tag{4.49}$$

This condition is necessary only and is naturally verified by the way the scheme is constructed. Indeed the numerical quadratures are based on the Taylor polynomial expansions of the response.

As a sample illustration for the Euler scheme from section 4.1.2, the limit can be computed using the quadrature of equation (4.4)

$$\lim_{h\to 0} \frac{q_{n+1} - q_n}{h} = \lim_{h\to 0} (1-\theta)\dot{q}_n + \theta q_{n+1} = \dot{q}_n \tag{4.50}$$

The same can be performed using equation (4.37) for the Newmark scheme.

The consistency feature is rather important as it provides an easy way to check the scheme solution. The optimal time step h to use is then the largest one for which the response computed does not vary.

The scheme convergence itself is based on the contracting properties of the iteration operator, usually not exhibited in the schemes formulation. The iteration operator $E(h)$ is obtained by writing the canonical form of evolution problems

$$u_{n+1} = [E(h)]\, u_n + f_{n+1}(h) \tag{4.51}$$

In the following, only the Newmark scheme is studied as it is the scheme that will be practically used. The classical methods featured in this section can however be applied to any time integration scheme. For the Newmark scheme detailed in equation (4.40), it can be proven [139] that

$$[E(h)] = \begin{bmatrix} M+\gamma hC & \gamma hK \\ \beta h^2 C & M+\beta h^2 K \end{bmatrix}^{-1} \begin{bmatrix} (1-\gamma)hC & (1-\gamma)hK \\ (\frac{1}{2}-\beta)h^2 C - hM & (\frac{1}{2}-\beta)h^2 K - M \end{bmatrix} \tag{4.52}$$

The stability of the scheme is obtained if the spectral radius of the evolution matrix is less than 1. Indeed, the response at time t_n is then expressed as a power of E, which is easily computed with its eigenvalues. If the spectral radius is greater than 1, then a perturbation of the response at time step n will grow indefinitively as function of time, generating divergence.

The evolution matrix *a priori* depends on the system considered, the time step and the scheme constants. In the case where the system is undamped, or if damping is low enough to consider that mode coupling does not occur, some results can however be obtained.

Assuming no damping occurs, the evolution matrix can be projected on the model real modes. For a given mode j, the iteration matrix can be computed explicitly and becomes

$$[E(h)] = \begin{bmatrix} 1-\gamma\frac{\omega_j^2 h^2}{1+\beta h^2\omega_j^2} & -\omega_j^2 h\left(1-\frac{\gamma\omega_j^2 h^2}{2(1+\beta\omega_j^2 h^2)}\right) \\ \frac{h}{1+\beta\omega_j^2 h^2} & 1-\frac{\omega_j^2 h^2}{2(1+\beta\omega_j^2 h^2)} \end{bmatrix} \tag{4.53}$$

The evolution matrix eigenvalues can then be computed, giving conditions on the time step and scheme coefficients for a given frequency. A solution under the form of complex conjugate eigenvalues exists for

$$(\gamma + \frac{1}{2})^2 - 4\beta \leq \frac{4}{\omega_j^2 h^2} \tag{4.54}$$

Writing the eigenvalues under the form $\rho e^{\pm i\phi}$ allows to evaluate the stability condition through the introduction of the value ξ, defined by

$$\xi^2 = \frac{\omega_j^2 h^2}{1 + \beta \omega_j^2 h^2} \tag{4.55}$$

lightening the expression of the eigenvalues modulus

$$\rho = \sqrt{1 - (\gamma - \frac{1}{2})\xi^2} \tag{4.56}$$

and angle

$$\phi = \arctan\left(\frac{\xi\sqrt{1 - \frac{1}{4}(\gamma + \frac{1}{2}^2\xi^2)}}{1 - \frac{1}{2}(\gamma + \frac{1}{2})\xi^2}\right) \tag{4.57}$$

Equation (4.56) then shows that the scheme is stable for $\gamma \geq \frac{1}{2}$. From equation (4.54), it can be noticed that the time step does not influence stability if $\beta \geq \frac{1}{4}(\gamma + \frac{1}{2})$. For such conditions, the scheme is then said unconditionally stable as the time step has no influence. These conditions are usually parametered by α so that the best coefficient choice for the Newmark scheme can be expressed as

$$\begin{cases} \gamma \geq \frac{1}{2} + \alpha \\ \beta \geq \frac{1}{4}(\gamma + \frac{1}{2})^2 \end{cases} \tag{4.58}$$

The stability conditions are traditionally reported on a 2D diagram parametered by β and γ, presented in figure 4.4

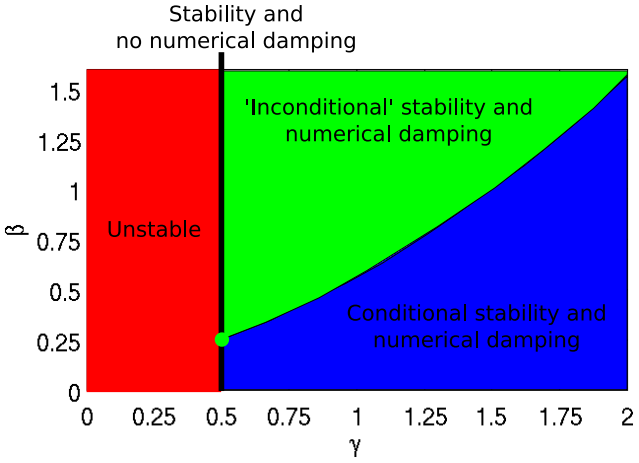


Figure 4.4: Newmark scheme stability diagram

The demonstration shows that unconditional stability can only be proven when modal decoupling can be performed, on linear models. For slightly damped structures, the expression of the scheme eigenvalues becomes more complicated, but a clear condition on low damping levels allow to conclude generally on the scheme stability.

The use of non linear models (which will generate a non linear evolution matrix), or with large damping or mode coupling do not elect for unconditional stability conditions. Detailed studies have been performed in the literature by Hughes [149] and later by Cardona and Géradin [150].

The fact that a scheme is consistent and stable does not suffice to provide a relevant scheme. The eigenvalues obtained in equations (4.56) and (4.57) for the Newmark scheme will indeed generate *amplification* and *periodicity* errors, which have a mechanical interpretation. In this scope, a time integration scheme can be viewed as a numerical filter, whose inner properties will impact the system response. Such study has for example been performed by Mugan [151].

The eigenvalues properties directly give the solution behavior. Equation (4.57) indicates that the simulation will show a phase evolution independent from the system. For classical implicit Newmark schemes, the periodicity error is proportional to $\omega_j^2 h^2$, which will be low for small type steps.

The amplification error has usually observable effects on the approximated response. Keeping the scheme eigenvalue complex form as $\rho e^{\pm i\phi}$ shows that ρ must be less or equal to 1. However, if $\rho < 1$, then the scheme response will be damped, independently from the system simulated. This phenomenon is classically known as *numerical damping*, and must be properly assessed. In particular the application is here targeting instabilities, which implies that the scheme to choose must allow proper energy increase patterns. Vola *et al.* [92] illustrate for example the effect of numerical damping in rubber wiper-glass squeal instabilities.

For the Newmark scheme, using the parameters defined in equation (4.58), the amplification error is expressed as

$$\rho - 1 = -\alpha \frac{\omega_j^2 h^2}{12} \tag{4.59}$$

The amplification error is then function of the time step and frequency. Using small time steps allows limiting the effect, but the fact that the amplification error will grow with the frequency is undesirable as rather high frequency instabilities will be important for brake squeal applications.

Attempts at controlling numerical damping in the Newmark scheme have yielded the Hilber-Hughes-Taylor scheme [136], keeping the Newmark numerical quadrature, with coefficients parametered by α from equation (4.58), but using a filtered residue signal.

$$\begin{aligned}[J]\,\Delta q_{n+1}^{k+1} = \quad & [M]\,\ddot{q}_{n+1}^{k} + \alpha\left([C]\,\dot{q}_{n+1}^{k} + [K]\,q_{n+1}^{k} - f_{cn+1} - f_{NLn+1}(q_{n+1}^{k}, \dot{q}_{n+1}^{k}, t_{n+1})\right) \\ & +(1-\alpha)\left([C]\,\dot{q}_n + [K]\,q_n - f_{cn} - f_{NLn}(q_n, \dot{q}_n, t_n)\right)\end{aligned} \tag{4.60}$$

This method is classically named HHT-α and features more interesting numerical damping values than the Newmark scheme. It has been studied in details in [150] which shows its exact properties. Its fundamental properties are however very close to the classical Newmark scheme, so that its interest is very limited in practice.

A global α-generalization was eventually proposed by Hulbert and Chung [152], and recently exploited for multibody dynamics with contact by Brüls and Arnold [153], where the inertial terms are also balanced by α.

From all existing finite differences schemes recent developments have converged toward the use of schemes of the Newmark family, with $\beta = \frac{1}{4}$ and $\gamma = \frac{1}{2}$ for most direct integration applications. The Newmark scheme does not introduce difficult to control numerical damping and features low periodicity errors, so that the scheme yielded shows very favorable stability conditions and remains quite neutral to the system.

For strong non-linearities, the linear evaluations do not stand. A direct application to a non-linear system must however be taken with care. In the case where Lagrange constraints are used the integration with second order schemes such as the Newmark becomes a problem.

The implementation of a constraint equation modifies the system equation (4.1), as it requires the use of Lagrange multipliers λ and a constraint matrix c,

$$\begin{cases} \begin{cases} M\ddot{q}(t) + C\dot{q}(t) + Kq(t) + f_c(t) + c^T\lambda(t) = 0 \\ cq(t) = 0 \end{cases} \\ q(0), \dot{q}(0), \ddot{q}(0) \\ \text{B.C. on } q, \dot{q} \end{cases} \tag{4.61}$$

Depending on the Newmark scheme formulation choice, the behavior for such a constrained system is different. If the system is undamped, and the scheme is acceleration based (linear and explicit applications), the system from equation (4.61) to solve at each time step is

$$\begin{bmatrix} M & c^T \\ c & 0 \end{bmatrix} \begin{Bmatrix} \ddot{q}_{n+1} \\ \lambda_{n+1} \end{Bmatrix} = \begin{Bmatrix} f_{n+1} - Kq_{n+1} \\ 0 \end{Bmatrix} \tag{4.62}$$

There, divergence occurs in the simulation due to the finite precision. Indeed, the second equation is always approximated regarding a tolerance ϵ, so that $c\ddot{q}_{n+1} = \epsilon$. The second equation of (4.62) is a dynamic equation without stiffness or damping to control the acceleration variations so that the imprecision yielded at a given time step will be integrated over time and thus grow indefinitely.

Baumgarte [89] gives a solution to such problem by suggesting to add a fake dynamic to the second equation of system (4.62). For a constant constraint, one can for example use

$$\begin{bmatrix} M & c^T \\ c & 0 \end{bmatrix} \begin{Bmatrix} \ddot{q}_{n+1} \\ \lambda_{n+1} \end{Bmatrix} = \begin{Bmatrix} f_{n+1} - Kq_{n+1} \\ -\alpha_b^2 q_{n+1} \end{Bmatrix} \tag{4.63}$$

A stiffness parameter α_b can be applied close to the system highest frequencies, the overall dynamic is then not disturbed while the constraint is verified and approximations are controlled. The same idea can be applied if the constraint is non linear.

Non linear implementations for an implicit scheme are displacement based with an iterative correction, so that the system to solve with a constraint becomes

$$\begin{bmatrix} J & c^T \\ c & 0 \end{bmatrix} \begin{Bmatrix} \Delta q_{n+1}^{k+1} \\ \Delta\lambda_{n+1}^{k+1} \end{Bmatrix} = \begin{Bmatrix} f_{n+1} - Kq_{n+1}^k \\ \frac{cq_{n+1}^k}{\beta h^2} \end{Bmatrix} \tag{4.64}$$

The behavior of system (4.64) is very different from system (4.62), as the constraint now constitutes a dynamic equation with stiffness but no inertia. The constraint behavior shows then infinite frequencies which cannot be integrated properly by the integration scheme. The α-generalized method of Hulbert and Chung [152] provides a way to numerically damp such behavior.

Constraints thus cannot be handled by an implicit non-linear Newmark scheme due to numerical instabilities. Lagrange contact is impossible to treat with such scheme.

4.1.6 Effect of the Newmark approximations in non-linear dynamics, illustration

The approximations generated by the time discretization can raise issues when dealing with non linearities. Indeed, non linear laws depending on the displacement or velocity will be interpolated by the simulation at computation points, so that the actual energy path the continuous system should have followed can be overestimated or underestimated. This is well exposed in [149] which features classical examples based on non linear springs. Such system is a common benchmark for the evaluation of time integration scheme for non linear systems, as presented in [154].

As an illustration, a simple non linear spring-mass system is considered. The system is of mass $m = 1kg$ and a non-linear spring is designed, of exponential stiffness k_{nl}, so that depending on the displacement q, the exponential tension force is expressed as

$$f_{nl}(q) = e^{\lambda q} \tag{4.65}$$

An exponential spring is considered in correlation with the contact laws that will be used in practice, for the system featured a value of $\lambda = 10m^{-1}$ is chosen.

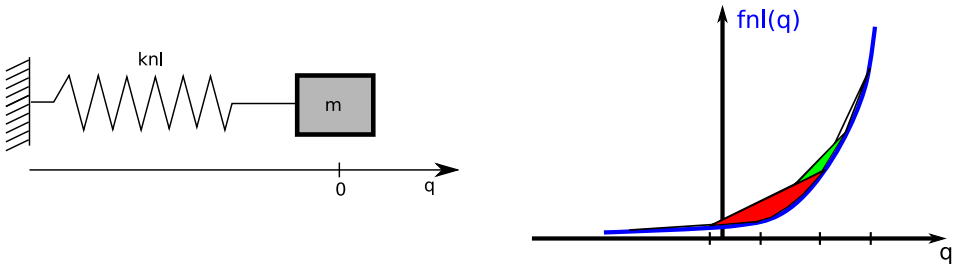


Figure 4.5: Exponential spring example showing approximations generated by the discretization in non linear cases, for a mass-spring system

Depending on the time discretization, the non linear stiffness evolution will be approximated, so that during an oscillation of sufficient amplitude, the non linear spring force will be overestimated, or underestimated, which is represented by the colored areas under the spring law curve of figure 4.5.

As function of the time step, the approximations can become very large. It is impossible to obtain exact energy compensation during a cycle, at least for numerical reasons for the spring example, but also because several non linear local behaviors can occur in a system like an automotive brake squeal.

In the case of a non linear mass-spring system, the equation of movement can be expressed as

$$\begin{cases} m\ddot{q}(t) + kq(t) + k_{nl}(q(t))q(t) = 0 \\ \\ q(0) = q_0 \ ; \ \dot{q}(0) = \dot{q}_0 \ ; \ \ddot{q}(0) = \ddot{q}_0 \end{cases} \tag{4.66}$$

Using in equation (4.66) $\ddot{q}_0 = 0$, $\dot{q}_0 = 25$ and $q_0 = 0$, for several time steps between $h = 10^{-4}s$, and $h = 2.5.10^{-2}s$ generates the results presented in figure 4.6, for a convergence determined once the adimensioned residue reaches $\epsilon = 10^{-12}$.

Due to the initial velocity, the mass goes first forward while decelerating exponentially, until the velocity becomes zero and the mass springs back. The linear stiffness is much smaller so that the spring compression phase is much longer than the spring elongation phase.

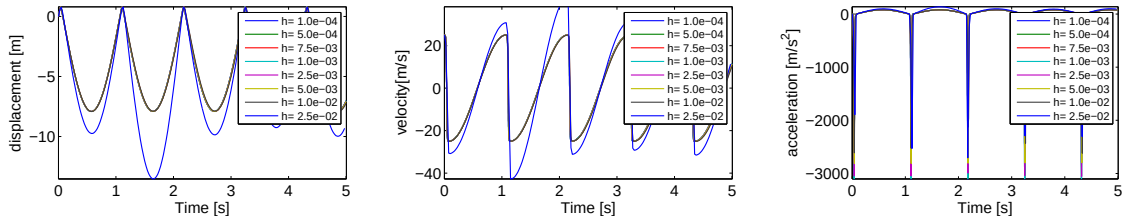


Figure 4.6: Time simulation of a mass exponential spring system, displacement, velocity and acceleration results for an initial velocity configuration

Depending on the time step chosen, results vary greatly. Figure 4.6 shows in particular that for a time step of $2.510^{-2}s$, the results are very poor. This clearly illustrates the fact that the choice of a time step for a system with strong non-linearities do not depend on mathematical convergence of the interpolation scheme but rather on the non-linearity scale. Indeed, the oscillation frequency yielded here is at $0.94Hz$, a bold application of classical rules for the choice of the time step would however indicate stability for every time step, thus a choice of $h = T/6 \simeq 1.7.10^{-1}s$ should have sufficed.

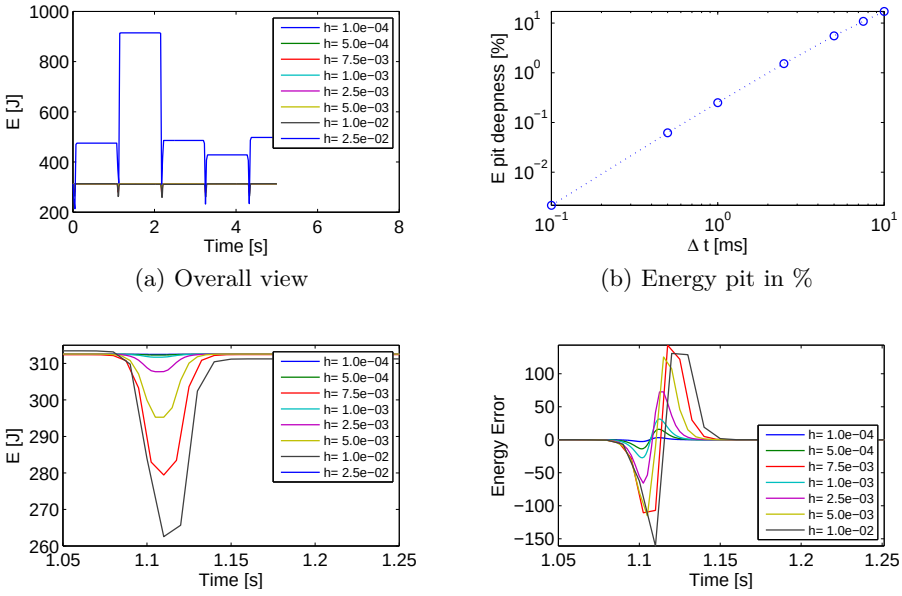
A posteriori however, the non-linear oscillation frequency should feature min/max values, from the face that the transitions at high spring compression are much quicker than the oscillation. From the min/max values of displacement, minimum and maximum stiffness values can be computed, from which a frequency interval is obtained. For the simulation results of figure 4.6, $F = [f_0; f_{max}] = [0.5; 28]Hz$. The *non-linear* time step to use should then be closer to $h = \frac{1}{6f_{max}} \simeq 6.10^{-3}s$.

Imprecisions occur mostly when the non-linearity is the strongest, here at the maximum elongation. As the exponential force increases, a small imprecision on the displacement will generate a large imprecision on the spring tension. This can be observed in figure 4.6 mostly on the acceleration signal.

The pattern is confirmed in figure 4.7, which plots the system mechanical energy. As the system is conservative, a constant mechanical energy is expected, which is not the case numerically. Figure 4.7a clearly shows that the large imprecisions for the largest time step tested occurs at the maximum elongation phase. The imprecision is here so large that the mechanical energy is poorly estimated, once the spring relaxes, it remains constant again, but at wrong levels.

Figure 4.7c shows in fact the existence of mechanical energy *pits* at the maximum elongation, whose deepness depends on the time step. Of course the finer the step the most precise the maximum elongation phase. The pits deepness are evaluated in figure 4.7b, as function of the time step.

Figure 4.7d illustrates the energy error generated by the numerical computation of the non-linear force work integral. Instead of integrating the exact energy from instant zero as in figure 4.7c, the work integral is performed assuming a linear variation between t_n and t_{n+1}, thus reproducing the approximations illustrated in figure 4.5. In the elongation phase, the spring work is resistive and is overestimated as it corresponds to an *ascending* phase, thus the mechanical energy decrease. Once the spring comes back from its maximum elongation, the spring work becomes positive and is once again overestimated (*descending* phase) so that the mechanical energy increases back. Hence the presence of energy pits.



(a) Overall view

(b) Energy pit in %

(c) Zoom in at 1s

(d) Non linear Work integral error

Figure 4.7: Mechanical energy of a mass-exponential spring launched with an initial velocity as function of the time step. Energy pits in percentage relative to the expected mechanical energy. Illustration of the integral approximation error (d), by computing the difference between the work assuming a piece-wise linear variation of the non linear exponential force between computation points and its exact formulation

The existence of these energy pits is a problem, as the energy restituted in a *descending* phase will not be the exact compensation of the loss in the *ascending* phase. It will affect the energy restituted to the system in the other phases, and is observed by the variations of mechanical energy levels in the constant phases where the non linearity is negligible (compression phase). The poor simulation using a time step of $h = 2.5.10^{-2}s$ illustrates in particular between figures 4.6 and figure 4.7a that the oscillation amplitude depends on the value obtained at the maximum elongation computed.

Although the simulations with a smaller time step seem to give relevant results, the energy pit pattern is still present, only at a smaller scale, which will raise issues for longer simulations. As the oscillation amplitude varies with the maximum elongation computed, the response phase may shift for long simulations. For the test case presented, after 50 oscillations, a clear shift appears depending on the time step, and plotted in figure 4.8.

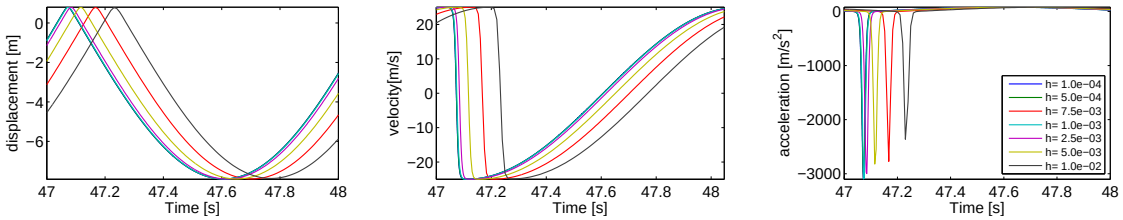


Figure 4.8: Displacement, velocity and acceleration signals for a 50s time simulation of the mass-exponential spring system

Between $h = 10^{-4}s$ and $h = 10^{-2}s$, the maximum elongation at time 47s is attained with a delay of 0.17s, which makes it a 0.3% periodicity error. The periodicity error estimated for $h = 10^{-2}s$, would be of $\frac{\omega^2 h^2}{12} = 2.9.10^{-2}\%$ for the non-linear frequency. Considering however a 28Hz oscillation based on the maximum instant frequency (largest tangent stiffness), the periodicity error for $h = 10^2 s$ becomes 25%. The scale of the non-linearity is then preponderant for such considerations.

It must be kept in mind that for a more complex system as a brake assembly, the non linearities can be scattered on different parts and/or be distributed. In such case, errors on non-linear transitions happen at different instants and the resulting interaction cannot be quantified as simply as for a mass-spring system.

The scale of the non linearity – whose *a priori* quantification is difficult – becomes then the only parameter to take into account to assess proper convergence properties, depending on the system simulated. Hughes [149] relates such considerations, by evaluating standard Newmark schemes, for linear and non-linear models. He proved in particular that under restricting conditions for the non-linearity (mostly on the continuity order), the Newmark scheme stability and convergence becomes conditional to the time step, but can be reached.

The interest of this section eventually resides in the fact that the non-linear law example is close to the contact laws provided by Bosch, at least in shape. The interaction with the chosen Newmark scheme is then interesting as some divergence patterns were observed.

4.2 Adaptation to a large industrial model

The time integration method choice is here very constrained by the thesis objectives. Indeed, it must be able to pass large models, for long simulations in a reasonable computation time. The section thus discusses choices made to obtain an acceptable strategy and select the associated parameters.

4.2.1 Rationale for the choice of a non-linear Newmark scheme

The interaction between the scheme and the non-linearity considered is critical, as detailed in sections 4.1.3 and 4.1.6. Signorini/Coulomb contact-friction laws generate mathematically difficult to handle signals, as acceleration can become discontinuous (*e.g.* due to a stick/slip or a slip-separation pattern).

The discontinuity itself will be impossible to reproduce numerically due to the time discretization. Specific time stepping methods, as discussed in [136], catch singularities better by estimating the instant at which a discontinuity occurs and by appending a specific time step. This is however unthinkable for an industrial system featuring several thousand contact DOF. Detecting status change occurrence for each point would slow down computation times dramatically.

The Non Smooth Contact Dynamics method [87] handles Lagrange contact-friction constrained with an adapted time integration scheme. Some adaptations are necessary, as the modified gap formulation presented in section 4.1.3. This formulation constitutes regularization as the contact constraint is taken into account before impact, based on the gap closing velocity. Decrease in the contact force is also associated with a non zero gap, so that augmentation parameters and convergence criteria play a significant role in the actual behavior. While this approach does seem promising, the stability of the iteration loop is difficult to handle and the topic is still studied in extension of the work of Lorang [6].

Practically, an effect similar to bilateral non linear contact laws (power or exponential type for example) should be observed for small open contacts.

Kenhous [138] presented a method to deal with second order time integration schemes and Lagrange contact formulations. The so-called redistributed mass method featured allows removal of irregularities for the impact of 3D solids discretized in finite elements. The solution suggested is to remove inertia from the contacting points, to obtain a discretized interface behavior closer to a continuous formulation. The mass terms associated to the contact DOF are then redistributed for each element on the non contacting points. This specific redistribution is performed using adequate shape functions for the mass matrix interpolation.

Formally speaking, the idea is to obtain an equation of the type

$$\begin{bmatrix} \tilde{M}_{cc} & 0 \\ 0 & 0 \end{bmatrix} \ddot{q}(t) + \begin{bmatrix} K_{cc} & K_{ci} \\ K_{ic} & K_{ii} \end{bmatrix} q(t) + c^T \lambda(t) = f_c(t) \tag{4.67}$$

where i denotes interface DOF, and c are complementary DOF. The method thus presented is equivalent to a penalization method, as pure stiffness terms are associated to the contact, following a stress-strain relationship.

Finally, the method proposed by Baillet *et al.* [8, 9] combines an explicit Newmark scheme with a Gauss Siedel resolution [85] for the Lagrange constraint. The regularization seems here minimal. The adjustments between the system displacement scale and contact scales are however performed at the following time step.

As for time finite element schemes, they still have to yield interesting results for non linear damped systems and they could thus not be considered.

Given these constraints, a functional representation of the contact pressure was retained as the only solution viable with certainty. This approach allows the use of standard schemes at the cost of losing some detail on impact and stick transition events. Furthermore, to avoid constraints on mesh quality an implicit solver is preferred. The non linear Newmark scheme, presented in section 4.1.4, was thus selected as a conservative strategy with adaptations discussed next.

4.2.2 Adapted Jacobian formulation

Selecting Newmark and a functional representation of contact pressure is not enough and the simulations still requires special treatment. From the 600,000 DOF model provided by Bosch, the reduction method suggested in section 2.2.5 leads to a 30,000 DOF model, which is much more acceptable but still intensive.

Indeed, if the non-linear Newmark scheme presented in figure 4.3 uses a non linear Jacobian formulation, this implies that in each Newton iteration of each time step, a Jacobian matrix is computed and factorized to resolve the displacement correction increment. For symmetric matrices of size 30,000 the most efficient solvers will require over 10s to factorize the Jacobian, which for 150,000 time steps will require at least 420 hours if only one resolution is necessary. A resulting 17 days computation is irrelevant for the application, where validation is sought for a target time of 8 to 12h.

The idea proposed to solve this difficulty is to use a fixed Jacobian, thus realizing a single factorization instead of several thousand. The feature presented in equation (4.48) is clearly mandatory.

Following section 3.2.4 notations, the derivative of the contact force with respect to state is given by

$$\frac{\partial f_N(q+\delta q)}{\partial \delta q} = [C_{NOR}]^T \left[\backslash \omega_j J(x_i) \backslash \right] \left[\frac{\partial p(q+\delta q)}{\partial \delta q} \right] [C_{NOR}] \tag{4.68}$$

where $\omega_j J(x_i)$ is the surface associated to each contact point, according to the notations of section 3.2.4.

Typically the derivative of the contact pressure $\frac{\partial p}{\partial \delta q}$ depends on the contact point and is a function of time. For a fixed Jacobian, the value must be set once and for all. In the present work, one further constrains the problem by using a stiffness density value k_c that is constant over the contact area. As a result, the proposed Jacobian takes the form

$$[J] = \frac{1}{\beta h^2}[M] + \frac{\gamma}{\beta h}[C] + [K] + k_c[\mathsf{K_c}] \tag{4.69}$$

where $[\mathsf{K_c}] = [C_{NOR}]^T [C_{NOR}]$.

It is useful to note that this choice is equivalent to considering the Jacobian of a system without non-linear forces but with a symmetric bilateral penalized contact at the interface.

The key point of the present discussion is that changing the Jacobian does not change the solution but may change convergence rates. k_c thus appears as a convergence parameter to take into account, along with the time step and tolerances. Its effect can however be studied in statics, as its influence is limited to the Newton-Raphson resolution loop.

Static convergence is studied for a range of k_c fixed Jacobian values. This is applied to a simplified brake example (presented in figure 1.13), and presented in figures 4.9 to 4.11. Static resolution for large displacements cannot be solved easily, as the fixed Jacobian limits the correction amplitude in the direction of contact. If the solids are not in contact and are far from each other, the long distance to correct will be done by increments inversely proportional to the Jacobian stiffness k_c. If contact is forecast, values in the hundreds of MPa will be sought for k_c, so that the maximum displacement increment will range in μm. If a gap of 1 cm must be closed, over 10,000 iterations will thus be needed, which is not acceptable.

For dynamic applications, or trivial static resolutions, the fixed Jacobian strategy is relevant, provided the small deformations assumption is valid. One considers the simple brake test case, with a constant pressure on the pads and an initial condition where the pads are physically in contact with zero interpenetration with the disc. Relevant stiffness density in the 50 and 500 MPa are tested in figure 4.9, which shows the number of iterations necessary to obtain convergence, or divergence if 200 iterations were performed without success. An optimum is clearly observed for which a minimal number of iterations is necessary.

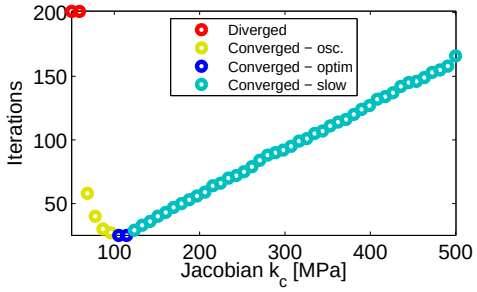


Figure 4.9: Number of iterations necessary to solve the simplified brake static state using the Newton-Raphson method and a fixed Jacobian k_c between 50 to 500 MPa. Convergence tolerance is fixed at $\epsilon_{disp} = 10^{-9}$ in a maximum of 200 iterations

Figure 4.10 illustrates the convergence pattern on the quadratic norm of the displacement response. The lower stiffness value lead to an oscillation pattern commonly called the *bouncing effect*. As the Jacobian is not stiff enough, correction increments are too large, so that the resolution oscillates around the solution with a given amplitude related to k_c. If the amplitude is too large, divergence occurs. A case between convergence and divergence (numerically speaking) is here observed as although no divergence occurs, the response norm oscillates between two fixed values. After the optimum, convergence slows down, as the correction increments become smaller.

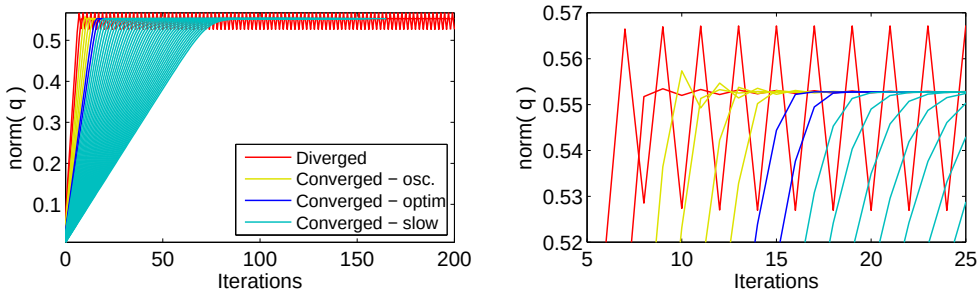


Figure 4.10: Convergence pattern plotting the displacement response quadratic norm as function of the iterations in the Newton-Raphson resolution and as function of the Jacobian parameter k_c. Left: global view. Right: zoom in a the optimized convergence points

Figure 4.11 illustrates the convergence pattern in a different way by plotting the quadratic norm of the displacement increment correction as function of the iteration and the stiffness k_c. The linear scale allows observing the first phase of the resolution when the displacement state is *far* from the equilibrium. The effect of the fixed Jacobian is here well illustrated by a constant norm depending on k_c, as already discussed.

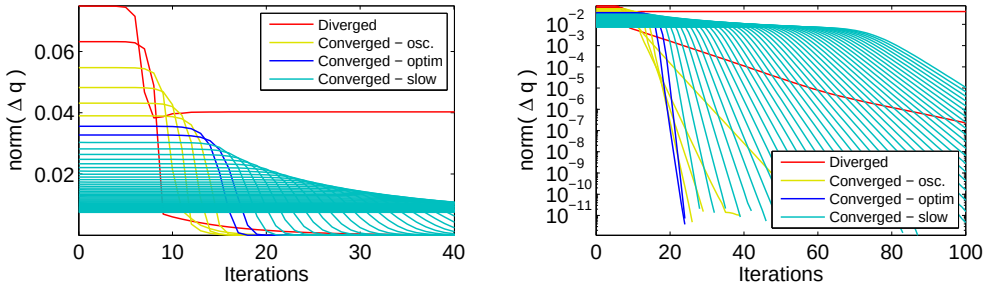
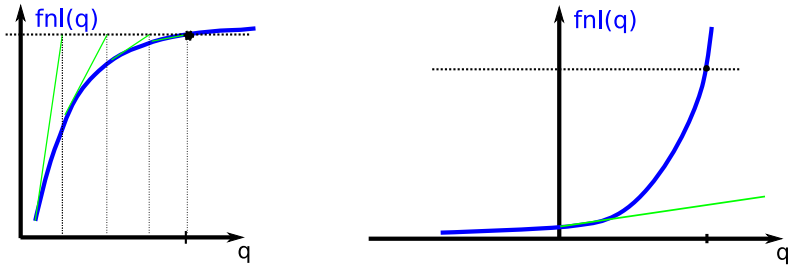


Figure 4.11: Convergence pattern plotting the correction increment quadratic norm as function of the iterations in the Newton-Raphson resolution and as function of the Jacobian parameter k_c. Left: linear y scale. Right: logarithmic y scale

Until contact happens, the smaller the Jacobian the better, indeed, the pad comes close to its final state quicker. The second phase of convergence is another story, as the increment corrections tends towards zero – this is better observed using a logarithmic scale. If the Jacobian is too flexible the smaller corrections are poorly performed such that oscillations or divergence occurs. If the Jacobian is too stiff the corrections are too small to converge quickly. The optimum is here characterized by the steepest slope observed.

From the algorithmic point of view, the Newton-Raphson procedure for the internal correction loop of figure 4.2 raises issues well illustrated in figure 4.12. The classical view of non-linear resolutions is illustrated with figure 4.12a, for which the resolution concept comes naturally.

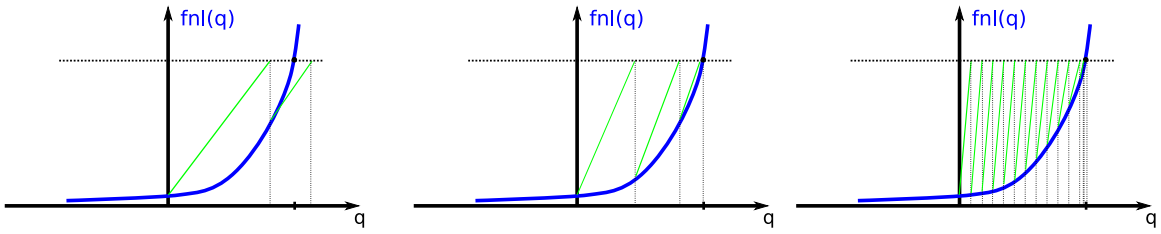
For stiffening problems, such as the type encountered with contact models when initial contact is set open or with very small interpenetration levels, the concept is less obvious. As shown in figure 4.12b, the first tangent stiffness can be very low so that the correction increment applied yields divergence.



(a) Softening problems (b) Stiffening problems

Figure 4.12: Convergence pattern for a genuine Newton-Raphson algorithm illustrated for a single DOF non linear spring depending on its evolution trend

The use of a fixed Jacobian in a stiffening configuration is illustrated by figure 4.13. For a low Jacobian stiffness, divergence occurs in figure 4.13a as the solution is passed over too quickly. If the Jacobian is overestimated in figure 4.13c, a large number of iterations is required, with a constant correction increment until the solution is close enough. An optimum thus exists for which the solution is approached with a convergence speed quick enough not to spend too much iteration and slow enough not to pass the solution point too quickly, figure 4.13b. The oscillations observed in figure 4.10 are then interpreted as a mild overshoot close to the solution.



(a) Too flexible : divergence (b) Optimized Jacobian (c) Too stiff : slow convergence

Figure 4.13: Convergence patterns for a Newton-Raphson algorithm with fixed Jacobian illustrated for a single DOF non linear spring showing an exponential shape

For problems difficult to solve with Newton iterations, control methods exist, as presented for example in [155, 90]. The idea is to evaluate the relevance of the Newton increment and to scale it to a realistic value. Such methods are based on minimization algorithms, and an efficient and computationally acceptable implementation is the use of a line search algorithm. This is a way of avoiding large iteration numbers when the fixed Jacobian is too stiff.

Conceptually, the equation to solve can be written

$$\{R(q)\} = [K(q)]\{q\} - \{f(q)\} = 0 \tag{4.70}$$

which is equivalent to the minimization of the quadratic form associated,

$$\min_{\{q\}} \left(\frac{1}{2} \{q\}^T [K(q)] \{q\} - \{q\}^T \{f(q)\} \right) \tag{4.71}$$

The Newton method computes correction increments with the help of the Jacobian

$$\{\Delta q\} = [J]^{-1} \{R\} \tag{4.72}$$

where the Jacobian is fixed in our case, and formulated with the help of $[\mathsf{K_c}]$, the contact matrix for a unit penalization,

$$[J] = [K_{el}] + k_c [\mathsf{K_c}] \tag{4.73}$$

The displacement increment $\{\Delta q\}$ is taken as a search direction, and an optimum increment along this direction $\rho\{\Delta q\}$ is searched. The secant method is employed to find the optimal coefficient ρ. The original minimization problem can then be written along the $\{\Delta q\}$ direction as

$$\min_{\rho\in\mathcal{R}}\left(\frac{1}{2}\{q+\rho\Delta q\}^T[K(q+\rho\Delta q)]\{q+\rho\Delta q\}-\{q\}^T\{f(q+\rho\Delta q)\}\right) \tag{4.74}$$

which can also be written

$$\{r(\rho)\}=\{\Delta q\}^T([K(q+\rho\Delta q)]\{q+\rho\Delta q\}-\{f(q+\rho\Delta q)\})=0 \tag{4.75}$$

or to keep the initial residue formulation of (4.70)

$$\{r(\rho)\}=\{\Delta q\}^T\{R(q+\rho\Delta q)\} \tag{4.76}$$

The secant method is a common resolution algorithm less efficient than the Newton algorithm, but that is also less costly since no matrix vector product is involved. The minimization of $r(\rho)$ is sought between two points ρ_i and ρ_{i+1} using the approximation of the tangent

$$\frac{dr}{d\rho}\sim\frac{r(\rho_{i+1})-r(\rho_i)}{\Delta\rho} \tag{4.77}$$

Solving $r(\rho)=0$ is then performed by a Newton method. With this particular tangent approximation, the iterative correction of ρ is then given by

$$\rho_{i+1}=\rho_i-\frac{\rho_i-\rho_{i-1}}{r_i-r_{i-1}}r_i \tag{4.78}$$

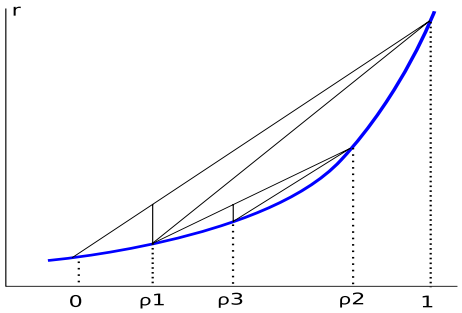
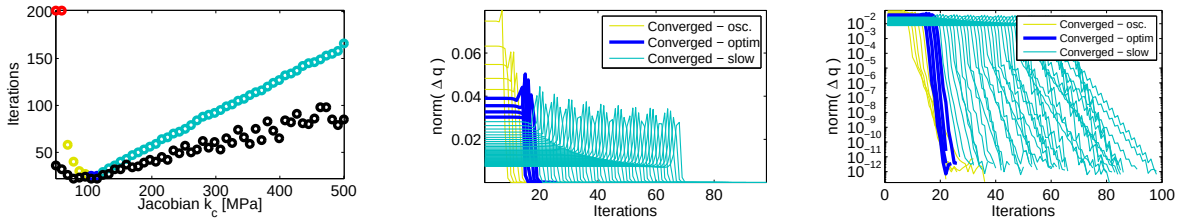


Figure 4.14: Illustration of the secant method for the line search algorithm

Only a few secant iterations are needed, since the optimum does not need to be accurate. In fact two secant iterations are enough to evaluate the allowable increment along the Newton increment direction.

This method works as a robustness enhancement of the Newton method as large increments due to high stiffness variations will be reduced while slow convergence due to high Jacobian values will be accelerated (with $\rho>1$).

To illustrate efficiency, the computations performed in figures 4.9 to 4.11 are now performed using the line search algorithm. Figure 4.15a shows an overall improvement, as no divergence occurs for the oscillating values, and as the number of necessary iterations is always reduced. It increases nevertheless when the Jacobian is not optimized and seems to increase linearly with the stiffness k_c once the optimum is passed. The oscillations of iteration number around the trend are explained by the fact that only two iterations are allowed for the line search, so that the optimum is only approximated.



(a) Number of iterations (b) Convergence pattern (lin scale) (c) Convergence pattern (log scale)

Figure 4.15: Improvement of the fixed Jacobian resolution by the use of a line search algorithm. From the plot of figure 4.9, the results for each points by the line search is added in black. Modified algorithm evolution of the number of iteration as function of the Jacobian

It is rather normal that the number of iterations increase when the Jacobian is too stiff due to the limitations of the line search algorithm. Indeed, if the initial correction increment provided by the fixed Jacobian is not optimized, the direction followed for the second optimization is not relevant, so that no improvement can be sought before convergence is close. This is illustrated in figure 4.15b, where the increment is still constant at the beginning.

Once convergence is close, the improvements of the line search algorithm can be observed. For too flexible Jacobians (low k_c) convergence is obtained as the increment norm is decreased dramatically. Comparing figures 4.13c and 4.15c illustrates the effect of the line search algorithm. For too stiff Jacobians, the increment corrections are increased by a factor 2 in the phase close to convergence, figure 4.15b.

In conclusion, the line search algorithm allows limited convergence improvements for contact problems. When the solution is far from convergence, no specific improvement is obtained in statics.

The method implementation requires additional residue computations, which is not wanted. For the target simulations, the cost of a residue computation is rather great, which makes this solution not fully relevant outside computation verification considerations. The solution retained for applications is a Jacobian optimization performed for once for a given project.

4.2.3 Numerical convergence strategy

In the perspective of section 4.1.6, convergence must be verified properly and independently from the mathematical assumptions coming from the numerical quadrature performed. Using regular contact-friction laws and the Newmark scheme, convergence is possible provided the time step is low enough [149]. In practice, an optimal time step can be found when the response variation is small (given a tolerance) if the time step is further decreased.

Two issues can be raised related to the computation of an optimal time step. First the tolerance to exit Newton-Raphson correction loops must be set. Second an indicator proving that the optimal time step is valid throughout the simulation is necessary.

Obtaining numerical convergence – *i.e.* $\Delta q \leq \epsilon_{comp} \simeq 10^{-16}$ – is generally counter-productive. The modified Newmark scheme with a fixed Jacobian does not provide quadratic convergence, so that convergence can be slow. Besides, the cost of a time simulation is the product of the number of time steps to perform and the number of iterations required in a time step, which requires an optimization. For a computational cost equivalent here, reducing the time step allows a refined response. It is thus preferable to reduce the step instead of over-iterating inside a time step.

The criterion used to check convergence varies depending on the problem to solve. The most reasonable option is to consider a *mechanical* convergence by taking the load residue as an indicator

$$\frac{\|r_n^k\|}{\|f_c\|} \leq \epsilon_f \tag{4.79}$$

As precision is difficult to obtain when computing the residue on large number differences, the convergence expected must be scaled against a reference value. The norm of the external forces is generally taken as the scale factor, but this can be discussed. It is in particular not possible for imposed displacement problems. A usual alternative is then to consider the residue of the first time step.

Alternatively, convergence can be checked through displacement increments

$$\frac{\|\Delta q_n^k\|}{\|q_n^k\|} \leq \epsilon_q \tag{4.80}$$

Again a scaling factor is preferable, taken as the current displacement response, if non null. In practice the existence and unicity of the solution is not fully guaranteed. The approximations generated at each level of the simulation makes this statement very difficult to interpret in practice. In particular, non-linear analyses, presented in section 3.3.3, that seek multiple solutions through continuation methods typically need more work in this area.

For a given convergence rate, the tolerance must be optimized. In practice, the same procedure and tolerance can be performed for time step and Newton iteration tolerance setting, that is to say no significant modification to the response is observed when the tolerance is decreased.

Relaxed convergence tolerances computed as detailed above are sufficient to obtain convergence, but do not fully protect against divergence over a long time simulation. Indeed, the system state can vary and the time step may become less relevant as function of time. The idea is then to check regularly throughout the simulation the convergence pattern. Every once in a while the algorithm is then forced to attain numerical convergence by over-iterating, if the convergence rate slows down a warning can be issued or the simulation stopped, depending on the convergence variation severity.

4.2.4 Details of importance

Further adaptation of the time scheme comes to nitty-gritty details having nevertheless a great impact on performance. First of all, the volume of data generated can be controlled by a sub-sampling strategy. As discussed in section 4.1.6, the non-linearity requires much smaller time steps than needed to represent the system dynamics itself. This results in very large response data-sets whose refinement is not always necessary. The idea is then to dynamically resample the response at a smaller frequency (coarser time discretization) by saving the response independently from the computation instants.

Practically, the Newmark quadrature presented in equation (4.36) allows interpolating the response for each instant between two time steps. The implementation suggested then uses two time discretizations, one dedicated to the actual time integration scheme operations and one dedicated to the instants to save. A factor 100 can exist between the time step needed for convergence and the one needed to properly express the system response, so that the amount of data output is reduced as well. A $100kHz$ subsampling is performed, although a $1GHz$ discretization at least is used for time integration.

Using subsampled signals allows very long time simulations and observation of the system response. Specific time intervals can be *a posteriori* found to be of interest and need to be observed in more details. A finer simulation for a small time interval based on the long simulation is then necessary. The procedure must therefore allow the simulation restart at any time.

This is rather simple to implement as the time integration scheme only requires a set of initial conditions on the displacement, velocity and acceleration. Such sets of data must be saved for a pre-determined set of restart points.

The most difficult part of optimization concerns the optimization of the algebraic operations needed by the time integration method. Observing the algorithm presented in figure 4.3, two major operations are present: the residue computation and the displacement increment resolution. These are two different applications as the first one consists in matrix vector products and interpolation of non linear forces computation, while the latter one is a linear resolution with a matrix inversion.

The sparse matrix vector products are efficient in MATLAB, but can be further optimized by developing dedicated procedures calling optimized algebraic computation libraries. The choice of the MKL, for Math Kernel Library, developed by Intel, has been done for SDT, through the use of MATLAB mex files. A substantial gain of 30% computation time was observed for the models tested here.

The resolution phase is already well optimized by the fact that the Jacobian is fixed. Indeed, it can be factorized once for all so that the resolution phase is already quick. The way the factorization is performed, depending on the matrix partitioning, as detailed in section 2.4.1 has however an impact on the resolution performance. An optimum set of factorization parameters exists for which resolution time can be reduced from the default setting.

4.3 Real brake time simulations

The time simulation method presented in section 4.2 is applied to the reduced model of section 2.2.5. Convergence is first assessed on complete simulations, before limit cycles are illustrated.

4.3.1 Convergence

The reduced model is fully converged in the frequency band of interest. After the general studies of the Newmark scheme of section 4.1.5, the choice of a relevant time step remains to be done. It must be kept in mind that the computation performance is balanced between the number of time steps to perform and the number of correction iterations inside a given time step.

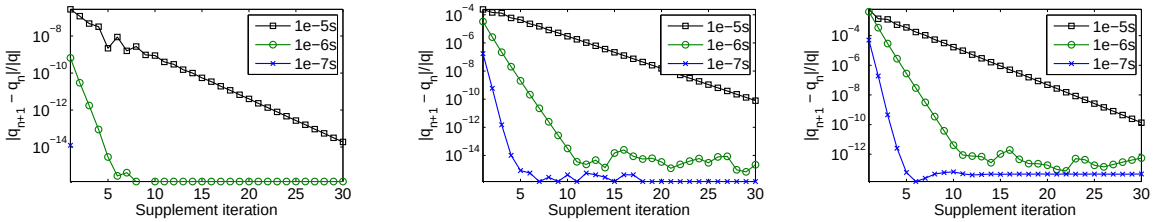
The algorithm convergence criterion is set as explained in section 4.2.3. The relevance of the chosen convergence criterion is here analyzed by forcing Newton-Raphson iterations after the algorithm convergence criterion is met.

The extended convergence patterns are here taken as the correction magnitude regarding the final value. Forcing the Newton-Raphson iterations, it is possible for each iteration k to compute *a posteriori* the evolution at time step $n+1$

$$\frac{\|q_{n+1}^{k+1} - q_{n+1}^{k}\|}{q_{n+1}} \tag{4.81}$$

Figure 4.16 illustrates the time step convergence behavior within a time step for three time step values. It has several influences, on the approximation generated by the prediction and on the Jacobian contracting properties. As the time step increases, the displacement variation to integrate is greater, requiring more iteration to obtain convergence.

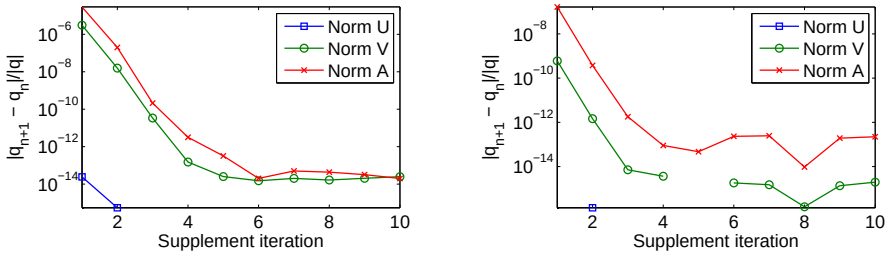
At 12 Bar, the $10^{-6}s$ time step seems correct regarding the initial levels, although numerical convergence is difficult to attain. The $10^{-5}s$ time step fails to converge in 30 iterations, which makes it irrelevant for the wanted simulations.



(a) Displacement (b) Velocity (c) Acceleration

Figure 4.16: Real brake time simulation convergence pattern as function of the time step at 12 Bar

Figure 4.17 shows that the convergence depends on the contact force magnitude, as the convergence pattern depends on the braking pressure. Indeed, for the same time step, numerical convergence is reached later at 30 Bar than at 12 Bar.



(a) 30 Bar simulation– $h = 10^{-7}s$ (b) 12 Bar simulation – $h = 10^{-7}s$

Figure 4.17: Real brake time simulation convergence pattern as function of the working point. The convergence criterion (4.81) is sequentially performed for U, V and A, for both time steps

The effect of a time step large enough not to yield a converged solution but without any divergence is analyzed in figure 4.18. The difference is clear between this and a refined time step solution.

The amplitude obtained is larger for the non converged case which shows a smoother solution. The non linearities integration is poor due to the large time step which makes the system respond in a different way, as previously presented in section 4.1.6. The FFT from figure 4.18 shows the response shows globally higher levels, especially for the peak at 8.5kHz.

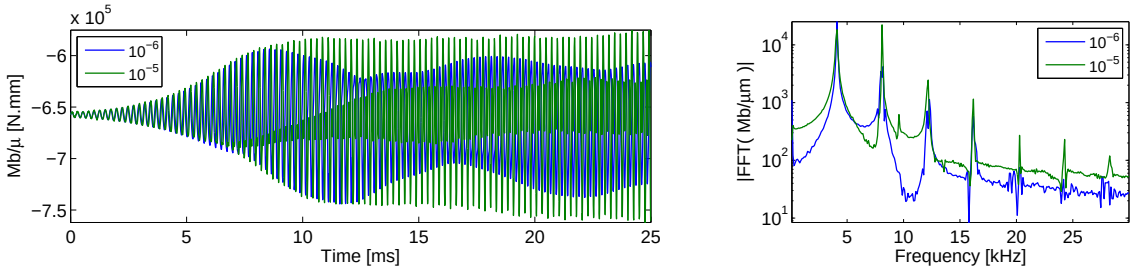


Figure 4.18: Braking torque signal obtained for the real brake model as function of the time step chosen. The FFT is computed from 10 to 20 ms

Displacements are also overestimated, as plotted in figure 4.19. Large differences (over 50%) can be observed which significantly alter the time deformation.

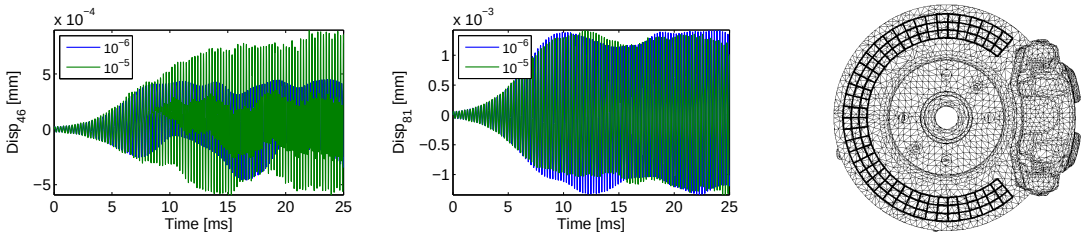


Figure 4.19: Displacement results for sample points of the ODS grid, as function of the time step

Variations of the total mechanical energy are another classical criterion of convergence. Figure 4.20 illustrates that larger times steps lead to significant errors after the initial growth phase.

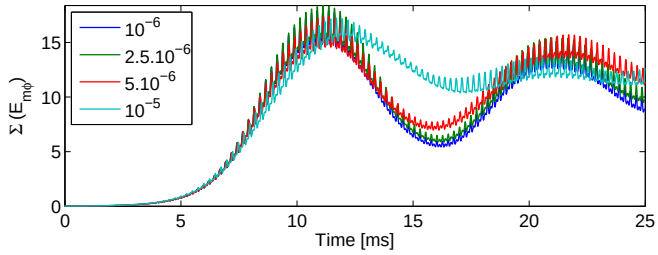


Figure 4.20: Total modal mechanical energy (modes R1 to R500) of the system as function of the time step

Based on the results shown, simulations performed at $10^{-6}s$ are considered as converged simulations in the basic case. The relative braking torque signal error is less than 0.1% throughout the simulation against the $10^{-7}s$ solution, as plotted in figure 4.21.

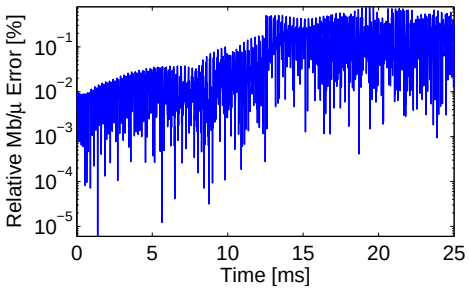


Figure 4.21: Relative braking torque error between simulations for $10^{-7}s$ and $10^{-6}s$ discretizations

The simulation error is less important at the beginning, which is in coherence with the observations that can be made with the displacements in figure 4.22. The displacement error increases with the simulation time more than the amplitude, which makes long simulations with time steps larger than $10^{-7}s$ conditionally acceptable.

The Newmark scheme introduces a time distortion (cf. section 4.1.5). It is proportional to the square of the time step so the effect in the simulations presented is small although not fully neglectable. As a perspective, convergence can be checked at several instants of the time simulation to ensure the convergence properties, tested by forcing Newton-Raphson iterations, are not altered over time. Restarts may also be performed throughout the simulation to reset the time distortion at several instants.

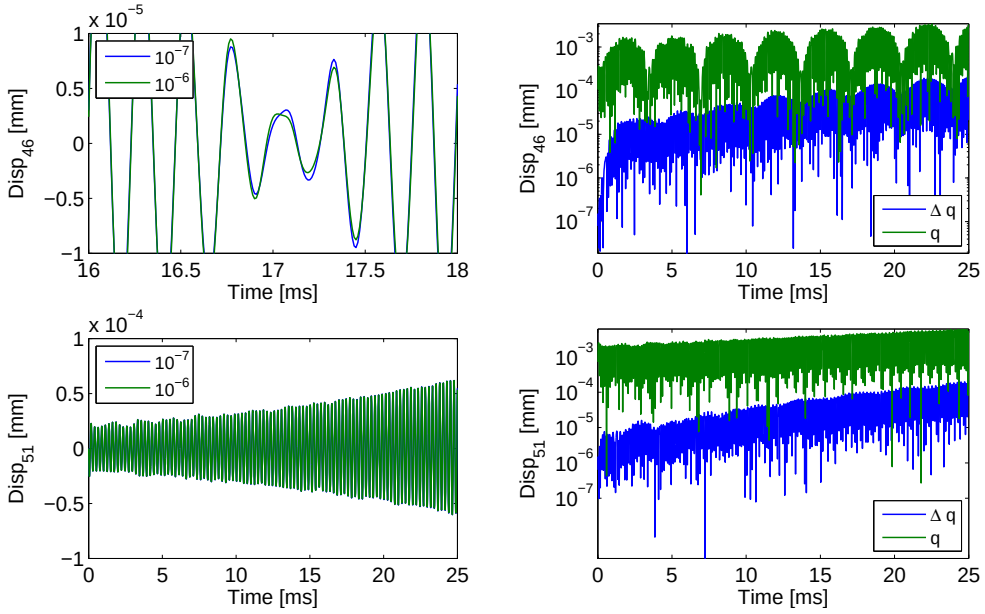


Figure 4.22: Relative displacement error between a $10^{-7}s$ and $10^{-6}s$ simulations featuring modal damping on the main instability. Left: displacement signals. Right: nominal signal at $10^{-7}s$ q and error between the $10^{-6}s$ and $10^{-7}s$ simulations Δq

Figure 4.23 eventually shows that the FFT quality decreases for time segments at the end of the simulation. In particular, great differences happen at the end of the 25 ms simulation, with spurious peaks in the working bandwidth (500Hz-16kHz). One can guess that higher amplitudes are the main source of this effect.

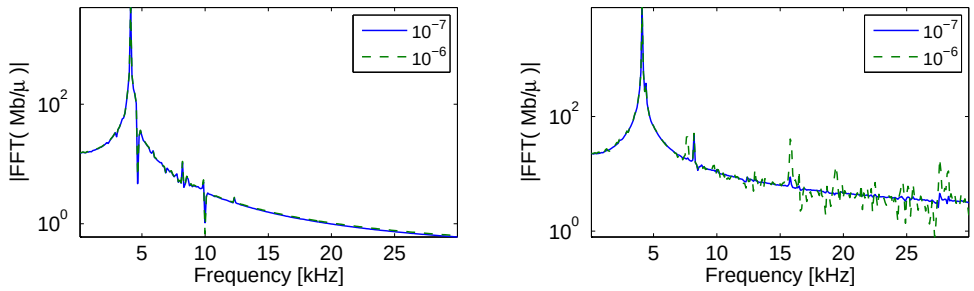


Figure 4.23: FFT of the braking torque signal between 1 and 10 ms (left) and 15 and 25 ms (right), as function of the time step

4.3.2 Limit cycle oscillations

In the procedure implemented, the static equilibrium is verified exactly, so that the Newmark residue (defined in equation (4.46)) is exactly equal to zero at the time origin. Some form of excitation is thus needed. Physical mechanisms are the variations of the rotating surface and variations of the applied pressure. Numerically the reproduction of physical mechanisms is not a necessity and strategies are discussed here.

In figure 4.24, motion is triggered by an over-pressure step of 10% during the first 200 ns of the simulation. This very short pressure step excites a wide bandwidth. As shown in figure 4.25a, high frequency modes are excited.

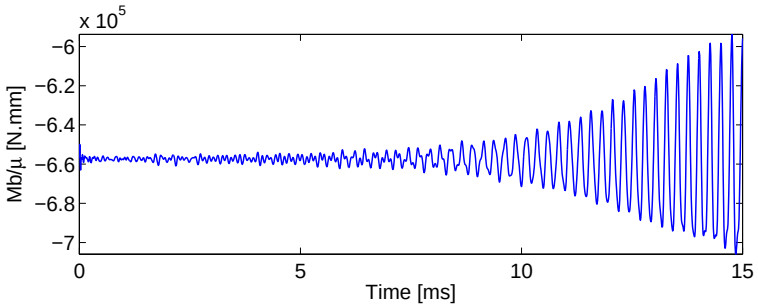
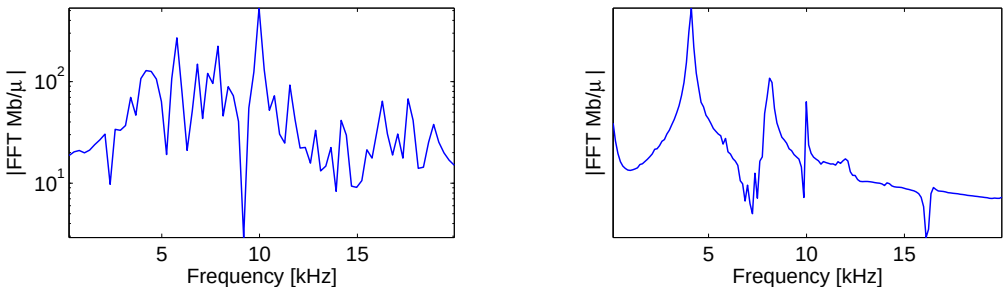


Figure 4.24: Squeal simulation triggered by a pressure step of 10% during 200ns

A transient evolution follows the initial oscillations after which a clear instability appears. The FFT of the signal end in figure 4.25b shows that the main resonance of the instability occurs at 4.2kHz. These results are in very good correlation with the experimental observations made by Bosch on this particular brake (section 1.2). A good confidence is thus placed in this working base.



(a) FFT from 0 to 4ms

(b) FFT from 7 to 15ms

Figure 4.25: Braking torque FFT during the simulation

It is interesting to note that the pressure step generates wave propagation along the disc, illustrated in figure 4.26. When the waves starting from each side of the pad meet, a global vibration is found which eventually settles. This observation is ascertained by the computation of a spectrogram plotted in figure 4.27. It can be seen that most modes are excited at the beginning of the simulation. A clear 4.2kHz response appears at 4ms, three harmonics then grow gradually.

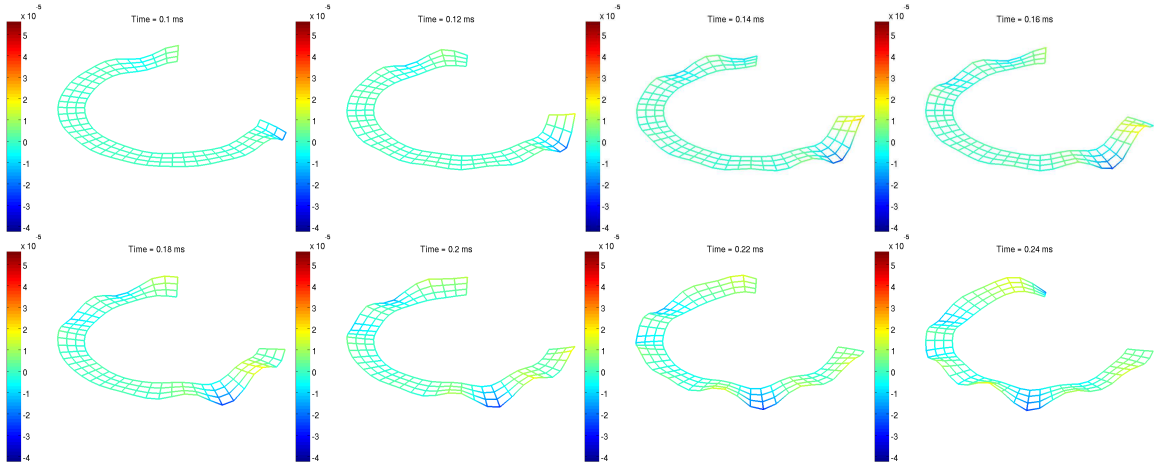


Figure 4.26: Vertical displacement response to the initial pressure step at the simulation start plotted on an observation mesh of the disc. Color code: displacement [mm]

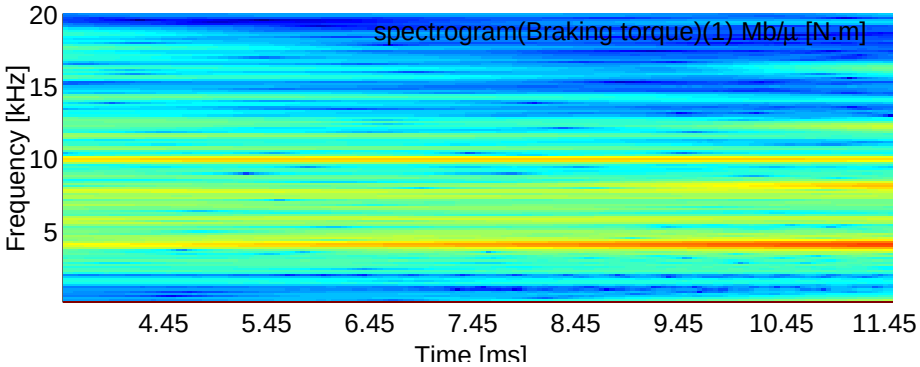


Figure 4.27: Braking torque spectrogram of the squeal simulation triggered by a pressure step

After the initial growth shown in figure 4.24, it can be seen that the signal shows a saturation pattern after which it is bounded. Such behavior is here called a *'limit cycle'*, although this is clearly not a proper mathematical characterization.

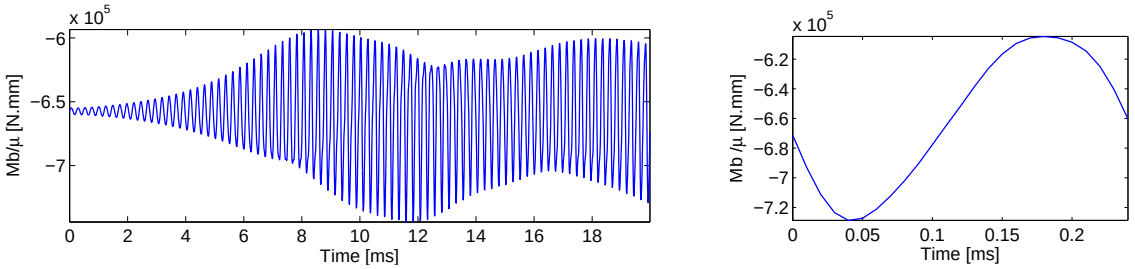


Figure 4.28: Limit cycle installation at 12 Bar triggered by mode C44 target and zoom in on a cycle at the simulation end

Figure 4.29 plots the vertical displacements of all sensors on the disc observation grid. Displacements up to $3\mu m$ can be observed. The distribution over half a cycle is plotted in the sequence figure 4.30, showing the space deformations. These displacement levels are coherent with the indications given by Bosch.

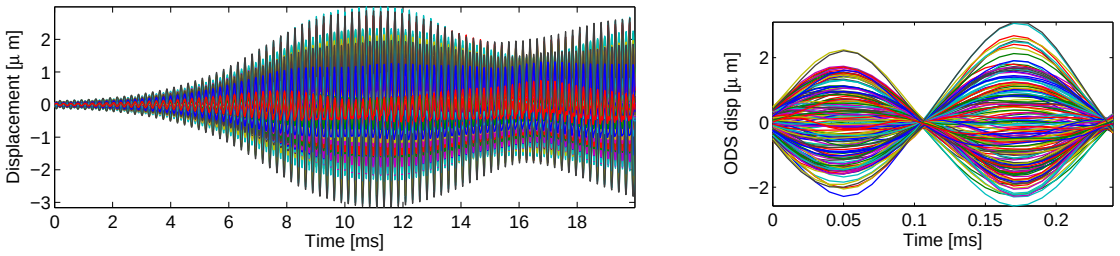


Figure 4.29: Limit cycle installation at 12 Bar trigged by mode C44 target, displacement measured on the disc ODS mesh, and zoom in on a cycle at the simulation end

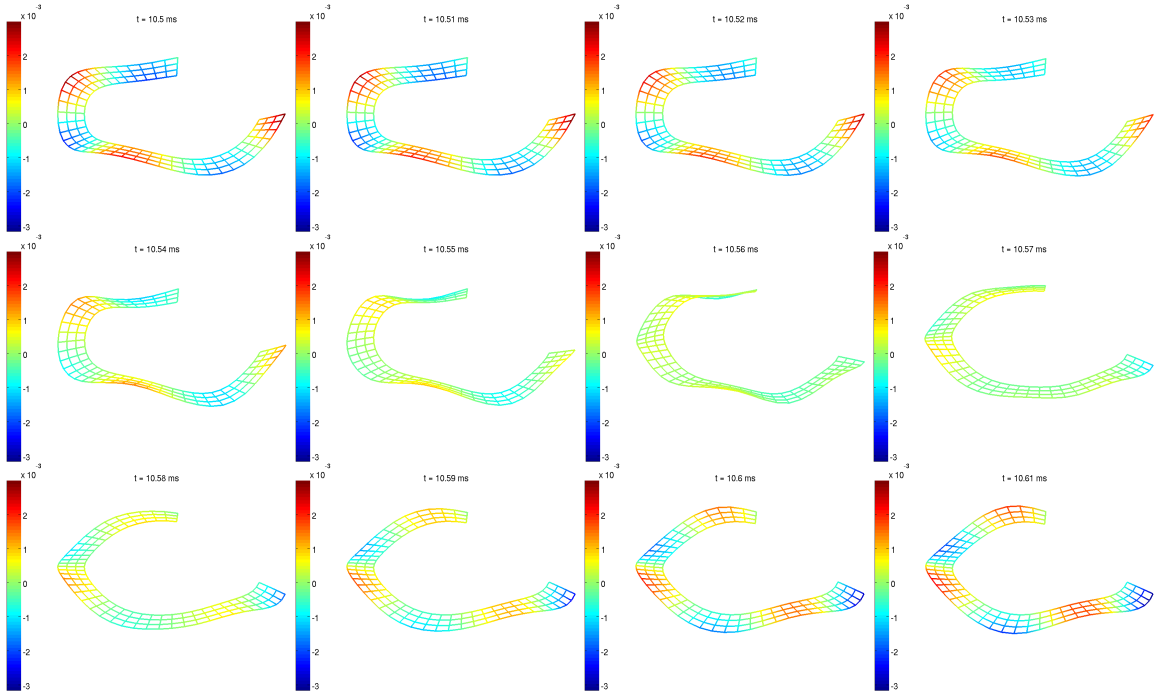


Figure 4.30: Limit cycle displacement plotted on the disc ODS mesh from 10.5 to 10.6ms. Color code: displacement [mm]

4.4 Design oriented time response exploitation

Classical design methods are based on modes. Structural optimization is performed by considering strain and kinetic energy distributions to give indications on how addition/removal of material may affect frequencies and possibly mode shapes. Damping modifications may also be interesting and a necessary step is then to identify modes that contribute to the overall response.

The end result of earlier developments is the demonstration of the ability to perform transient analysis of brakes. The use of spectrograms allows identifying frequencies by FFT and their evolution through time, but this information does not readily translate into information about modes that need to be modified.

To allow design iterations, the section first addresses damping models and introduces a new methodology to allow the introduction of modal damping models in the time simulations. The use of modal damping is very much related to the notion of modal amplitudes, so that modal sensors are introduced in section 4.4.3.

Finally, information about the energy distribution associated with the response in the stabilized phase may be as relevant as the analysis of contributing modes. Section 4.4.4 thus shows how a singular value decomposition of the limit cycle can be used to analyze contributions.

The tools presented are then used in section 4.5.

4.4.1 Common damping models

Realistic damping modeling is necessary to obtain relevant simulations, yet difficult to obtain. Common damping sources are friction at the interfaces and material dissipation, which is typically viscoelastic.

For viscoelastic damping, stress and strain are assumed to be related by the convolution with a relaxation function

$$\sigma(t) = \int_0^t E(t-\tau)\dot{\epsilon}(\tau)d\tau \tag{4.82}$$

The Laplace transform of the relaxation function is a complex modulus

$$E^*(\omega) = E_0(\omega)(1 + i\eta(\omega)) \tag{4.83}$$

which is frequency and temperature dependent. Figure 4.31 illustrates a typical nomogram describing these dependencies as described in the book of Nashif [156], the Viscoelastic Toolbox documentation [157], or by Coleman and Noll [158] treating of linear viscoelasticity. The dissipation in the material is related to the loss factor η which gives the dephasing between stress and strain. The use of complex moduli is also often called *hysteretic* damping but does not have a direct time domain equivalent.

The classical approach for time domain implementation is the definition of internal states featuring relaxation times. In this case, the viscoelastic material is represented by rheological models, featuring specific setups of springs and dashpots grouped as cells. The Kelvin cell is a basic spring-dashpot cell in parallel, which has been generalized in either Kelvin Voigt cells, where Kelvin cells are in series, or by Maxwell cells where serial spring-dashpot systems are put in parallel, as presented in figure 4.32.

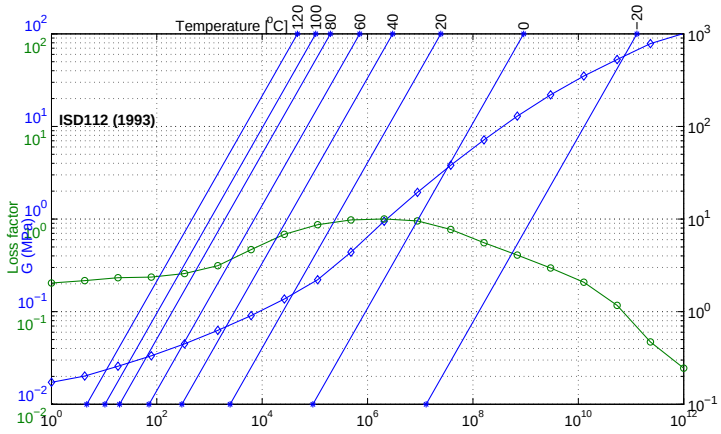
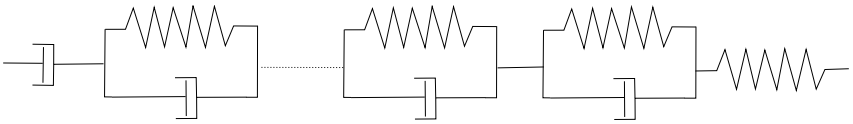
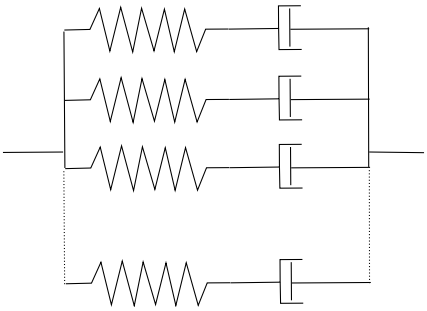


Figure 4.31: Nomogram of the ISD112



(a) Kelvin Voigt cells



(b) Generalized Maxwell cells

Figure 4.32: Viscoelastic rheological models, using cells of spring dashpot assemblies

The stress-strain relationship given by such cell systems can be written using differential operators,

$$\sum_{k=0}^{M} a_k \frac{d^k \sigma}{dt^k} = \sum_{l=0}^{N} b_l \frac{d^l \epsilon}{dt^l} \tag{4.84}$$

which can be generalized using fractional derivatives [159].

In the case of generalized Maxwell cells, the relaxation modulus given can be written using Prony series [160],

$$E(t) = E_0 + \sum_{l=0}^{M} E_l e^{\frac{t}{\tau_l}} \tag{4.85}$$

In the frequency domain, this yields the definition of hysteretic stiffness matrices of the type

$$K(\omega) = K_0 + \sum_{l=0}^{M} \frac{iK_l C_l \omega}{K_l + iC_l \omega} \tag{4.86}$$

where K_l and C_l are stiffness and damping properties of the l^{th} cell.

The aim of such models is to represent classical viscoelastic behavior using general parametric models. Park [161] gives for example a rather exhaustive presentation of such models. These formulations are great to use in the frequency domain as matrices can be computed and complex modes can be solved. In the time domain however, these formulations lead to higher order models, often requiring special purpose time integration mechanisms, which were ruled out from this work.

Another more general issue comes from the fact that damping is typically associated to matrices, which implies that the dissipation is realized by a modeled part. Finite element modeling however usually avoids precise junctions representations, although dissipation cannot be neglected at the interfaces.

The only solution deemed acceptable here in terms of computation times is thus a viscous damping matrix defined on the classical displacement DOF. Most applications use Rayleigh damping, where the viscous damping is considered as a linear combination of the mass and stiffness matrices, using mass coefficient α_R and stiffness coefficient β_R, on the assembled system,

$$[C_{Ra}] = \alpha_R [M] + \beta_R [K] \tag{4.87}$$

Rayleigh damping can also be performed piece-wise, *i.e.* by material weighting. For each material e, coefficients α_{Re} and β_{Re} are applied to the associated element matrices

$$[C_{Ra}] = \sum_e \alpha_{Re} [M_e] + \beta_{Re} [K_e] \tag{4.88}$$

Except in specific cases like aerodynamic forces or viscous dampers, the viscous damping matrix has no physical interpretation in the FEM coordinates. The usual relevant interpretation is related to the Modal Strain Energy method (MSE) [162], which assumes that the real modes of the undamped system are representative of the true motion and that viscous damping in modal coordinates represents an equivalent of the energy dissipated over a cycle. One thus assumes that a diagonal matrix in modal coordinates is physically representative

$$\left[\backslash 2\zeta_j\omega_j \backslash\right] \approx [\Phi]^T [C] [\Phi] \tag{4.89}$$

The MSE can be applied using most dissipation mechanisms: viscoelastic materials represented by a complex modulus, systems with friction, *etc.* For complex moduli, the classical interpretation is that the damping ratio ζ_j of a given real mode j will be the combination of the loss factor in each material weighted by energy contribution

$$2\zeta_j = \sum_e \eta_e \frac{\phi_{j|e}^T [K_e] \phi_{j|e}}{\phi_j^T [K] \phi_j} \tag{4.90}$$

where $\phi_{j|e}$ is the j^{th} real mode trace on the DOF corresponding to element e, $[K]$ the assembled stiffness matrix, $[K_e]$ the stiffness matrix associated to element e and η_e the loss factor of the material in element e.

Although in real life the modes computed with damping are really complex and therefore different from the real modes, the assumption works well, in particular for lightly damped structures, as originally discussed by Hasselman [163] or Caughey [164], and more recently exploited by Balmes [165], in the viscoelastic design tools [157].

Similar treatments can be performed for piecewise Rayleigh damping, or materials with a loss factor, which is the main interest of the MSE. Indeed, under these assumptions, for global Rayleigh damping (4.87), using real modes and their orthogonality conditions one has

$$\gamma_j = 2\zeta_j\omega_j = \phi_j^T (\alpha_R [M] + \beta_R [K]) \phi_j = \alpha_R + \beta_R \omega_j^2 \tag{4.91}$$

giving the classical modal damping ratio

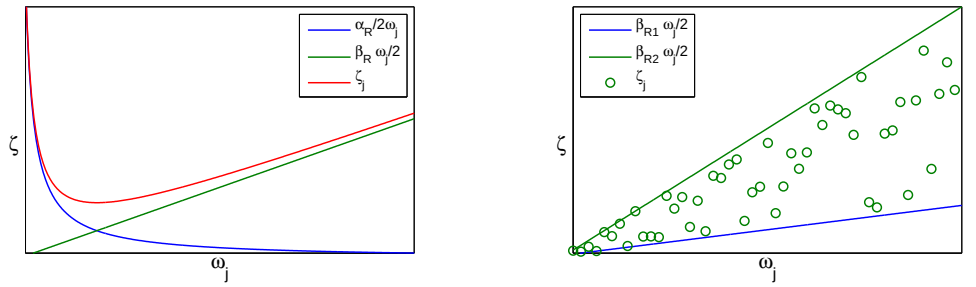
$$2\zeta_j = \frac{\alpha_R}{\omega_j} + \beta_R \omega_j \tag{4.92}$$

plotted in figure 4.33a.

For piecewise Rayleigh damping, the damping matrix of equation (4.88) projected on the real modes is not diagonal. The light damping assumptions consider however that outer diagonal terms are negligible, giving the modal damping ratios

$$\gamma_j = 2\zeta_j\omega_j = \sum_e \left(\alpha_e \left(\phi_{j|e}^T \left[M_e\right] \phi_{j|e}\right) + \beta_e \left(\phi_{j|e}^T \left[K_e\right] \phi_{j|e}\right)\right) \tag{4.93}$$

whose distribution pattern is specific. This is conceptually represented in figure 4.33b, considering only a piecewise Rayleigh damping affected to the stiffness matrix, for two distinct materials. Real life applications can be found in [166] by Bianchi *et al.* for railway applications – in pantograph-catenary interactions.



(a) Global Rayleigh damping

(b) Piecewise stiffness Rayleigh damping

Figure 4.33: Modal damping ratios obtained by the MSE method for Rayleigh damping formulations

Problems with Rayleigh damping are visible in figure 4.33. Damping associated to the mass matrix generates high ratios which is rarely representative of the first modes of a structure. Associating damping to the stiffness matrix features a damping ratio linearly increasing with the frequency, which will limit the maximum working frequency. The Rayleigh damping parameters must then be tuned for a given frequency band of interest whose width will be limited. Brake squeal applications target a wide working frequency band from $500Hz$ to $16kHz$ thus making difficult the setting of a set of Rayleigh parameters. Rayleigh damping is eventually not fully satisfying as hysteretic damping ratios will tend to be constant in frequency.

A global Rayleigh damping on the stiffness of $10^{-7}\,[K]$ was used in all the simulations presented, corresponding to a 1% damping ratio at $40kHz$.

4.4.2 Time domain modal damping

Under the modal damping assumption, one assumes the availability of modal damping ratios, coming from experimental characterization for example, for a set of low frequency modes. One thus has

$$\left[\backslash 2\zeta_{Mj}\omega_{j}\backslash\right]_{1\leq j\leq NM} = [\Phi]^T [C_{Mfe}] [\Phi] \tag{4.94}$$

This assumption is widely used to implement general damping, independently from materials or junctions. In the frequency and time domain, modal space solutions available in most finite element software packages. For the considered application, one however wants to perform time domain integration in the physical and not in the modal coordinates.

A recovery of the viscous damping matrix in physical coordinates is thus needed. Mass normalized modes verify

$$[\Phi]^T [M] [\Phi] = \mathcal{I} \tag{4.95}$$

which implies that the inverse of $[\Phi]$ is

$$[\Phi]^{-1} = [\Phi]^T [M] \tag{4.96}$$

The expression of the modal damping matrix in the finite element basis is thus given by

$$[C_{Mfe}] = [M] [\Phi] [\mathcal{C}_{\mathcal{M}}] [\Phi]^T [M] \tag{4.97}$$

The implementation of modal damping raises computational issues, as the finite element modal damping matrix obtained in equation (4.97) is full. This is obviously not possible to handle for large finite elements models such as the industrial brake.

The time domain modal damping matrix is thus stored implicitly, by noting that once the real modes are computed, and $[C_{other}]$ is projected, equation (4.97) can be written as a matrix-vector product sequence:

$$[C_{Mfe}] = \sum_j [M] \phi_j (2\zeta_{Mj}\omega_j) \phi_j^T [M] \tag{4.98}$$

In the implementation one stores the set of vectors $[M]\phi_j$ and the scalars $2\zeta_j\omega_j$ to compute sequentially the application of the damping matrix to the velocity vector $\dot{q}$, following three steps. First the time response projection $\phi_j^T [M] \dot{q}$, second, application of the scalar coefficient, third, computation of the load proportional to $[M]\phi_j$.

Modal truncation can be performed to improve computation times, noting that the first modes are typically less damped. Besides, modal damping outside the frequency band of interest ($500Hz$-$16kHz$) is not specifically relevant, so that a less costly Rayleigh damping can be used for these modes.

Implementing other damping sources must be carefully performed. One wants here to control the damping ratio of each mode, so that target values ζ_j can be applied explicitly. In combination with other viscous damping matrices (namely gathered in $[C_{ohter}]$), each mode damping ratio must be tuned to take into account the existing ratio coming from other damping sources. The modal damping ratio ζ_{Mj} to input in $[C_{Mfe}]$ must therefore be adapted to yield the targeted value ζ_j.

Assuming a damping matrix $[C_{other}]$ is associated to the model, implementing time modal damping in addition writes

$$[C_{fe}] = [C_{other}] + [C_{Mfe}] \tag{4.99}$$

such that the actual targeted modal damping ratio ζ_j characterizing the full system writes

$$2\zeta_j\omega_j = 2\zeta_{Mj}\omega_j + \phi_j^T [C_{other}] \phi_j \tag{4.100}$$

If Rayleigh damping is used in addition to modal damping, $[C_{other}] = [C_{Ra}] = \alpha_R [M] + \beta_R [K]$. The modal damping coefficients ζ_{Mj} to be input to obtain the wanted modal damping ratios ζ_j are then

$$2\zeta_{Mj}\omega_j = 2\zeta_j\omega_j - \alpha_R - \beta_R\omega_j^2 \tag{4.101}$$

4.4.3 Modal sensors

This time modal damping formulation (4.98) highlights the possibility to compute the projection of the time response on the modal basis, $\phi_j^T [M] \dot{q}$, which opens perspectives regarding modal identification from a time simulation, through the use of so-called modal sensors. Complex modal sensors were suggested by Lorang [6], which makes exploitation difficult in terms of energy. The application presented here uses real modal sensors based on modal mechanical energy, which is more easily handled.

The system response $\{q(t)\}$ is expressed as a linear combination of its real modes $[\Phi]_{N\times N}$, using coefficients $\{\alpha(t)\}_{N\times 1}$,

$$\{q(t)\} = [\Phi] \{\alpha(t)\} \tag{4.102}$$

Due to the mass normalization, the vector $\{\alpha(t)\}$ can be recovered, noting that

$$[M\Phi]^T \{q(t)\} = [M\Phi]^T [\Phi] \{\alpha(t)\} = \{\alpha(t)\} \tag{4.103}$$

which then gives the transient modal participation of each mode of the structure.

If the modal basis is truncated, keeping only NM modes, the displacement can be expressed as

$$\{q(t)\} = \sum_{j=1}^{NM} \{\phi_j\} \alpha_j(t) + \{q_R(t)\} \tag{4.104}$$

where the residual displacement $\{q_R(t)\}$ is orthogonal to the contribution of the first NM modes, as it verifies

$$\{q_R(t)\} = \sum_{j=NM+1}^{N} \{\phi_j\} \alpha_j \tag{4.105}$$

In practice, $\{q_R(t)\}$ is computed by subtracting the kept modal contributions from $\{q(t)\}$.

The same operation can be performed on the velocity $\{\dot{q}(t)\}$. This eventually gives the $\alpha_j(t)$ and $\dot{\alpha}_j(t)$ coefficients, using equation (4.103),

$$\begin{cases} \alpha_j(t) = \phi_j^T [M] q(t) \\ \\ \dot{\alpha}_j(t) = \phi_j^T [M] \dot{q}(t) \end{cases} \tag{4.106}$$

The modal amplitudes are varying over time so that a direct visualization is difficult to exploit. However, since the modes are orthogonal, the total energy is equal to the sum of modal energies. The modal displacements in physical coordinates are given by $\left\{q_{\phi_j}(t)\right\} = \alpha_j(t) \{\phi_j\}$ and the expression for velocity is similar. The modal kinetic and strain energies are thus given by

$$\begin{cases} 2E_{cj}(t) = \dot{\alpha}_j(t) \left(\phi_j^T [M] \phi_j\right) \dot{\alpha}_j(t) \\ \\ 2E_{pj}(t) = \alpha_j(t) \left(\phi_j^T [K] \phi_j\right) \alpha_j(t) \end{cases} \tag{4.107}$$

The total mechanical energy contribution of mode ϕ_j in the time response is thus

$$2E_{mj}(t) = \dot{\alpha}_j^2(t) + \alpha_j^2(t)\omega_j^2 \tag{4.108}$$

This energy is a good indicator of modes responding in the time simulation. The mechanical energy levels allow to observe variations and possibly interactions between modes (*e.g.* forced response), as well as the divergence ratios seen in the early phases of a limit cycle instability.

A full characterization of the system non linear response over its real modes can be flawed if a significant part of the system full mechanical energy is outside of the normal modes retained, or associated with non-linear loads not represented in the stiffness matrix.

For braking applications, friction may not only couple real modes in a certain way but also increase the space spanned by the system over time. The participation of $q_R(t)$ defined in equation (4.105) can become large enough not to be neglected. A simple way to assess it *a posteriori* is to compute the residual mechanical energy between the total mechanical energy and the modal one

$$2E_{mr}(t) = \{q(t)\}^T [K] \{q(t)\} + \{\dot{q}(t)\}^T [M] \{\dot{q}(t)\} - \sum_{j=1}^{NM} 2E_{mj}(t) \tag{4.109}$$

It can be noticed that the residual mechanical energy (4.109) corresponds to the residual displacement and velocity contributions $q_R(t)$ and $\dot{q}_R(t)$ defined in equation (4.105),

$$2E_{mr}(t) = \{q_R(t)\}^T [K] \{q_R(t)\} + \{\dot{q}_R(t)\}^T [M] \{\dot{q}_R(t)\} \tag{4.110}$$

System modes are usually computed around a steady state. Since the steady state is present when computing transient simulations, it is usually necessary to remove it from the time response $\{q(t)\}$ before modal projections,

$$\{\tilde{q}(t)\} = \{q(t)\} - \{q_0\} \tag{4.111}$$

the projections (4.106) should then be performed on $\{\tilde{q}(t)\}$.

4.4.4 Space-time decomposition of a limit cycle

The idea behind the use of modal sensors is to identify shapes (a space dimension) in the response, along with a level of contribution over time (a time dimension). With this objective, modal sensors target predefined shapes only (the system real modes), which may be too restraining for the analysis wanted. Using the Singular Value Decomposition (SVD) allows another space-time identification with no *a priori* shape targeting. For the decomposition to be applicable, it must be applied to a single or few cycles. This limits its application to established limit cycles.

The SVD is a classical generalization of diagonalization for rectangular matrices, details can be found for example in the book of Strang [167]. It can thus be used for subspace methods, in the way explained in this section.

Given a matrix $T_{N_H \times N_T}$ seen as an application from a space $\mathcal{R}^H$ to $\mathcal{R}^T$, the SVD computes $U_{N_H \times N_H}$, $\Sigma_{N_H \times N_T}$ and $V_{N_T \times N_T}$ so that

$$\begin{cases} [T][V] = [U][\Sigma] \\ [U]^T [U] = \mathcal{I} \\ [V]^T [V] = \mathcal{I} \end{cases} \tag{4.112}$$

U and V are then orthogonal and thus respectively represents a basis of $\mathcal{R}^T$ and $\mathcal{R}^H$. Σ is rectangular with a diagonal topology, the extra diagonal part (since it is a rectangular matrix) is filled with zeros. The diagonal terms are placed in decreasing order.

Figure 4.34 illustrates how the SVD analyses the effect of the linear operator T. An arbitrary vector is projected on a basis V of space $\mathcal{R}^T$. Each component is then mapped to the corresponding basis vectors of a basis U of space $\mathcal{R}^H$ and multiplied by the scaling factor $\sigma_i = [\Sigma_{ii}]_{i \leq \min(N_H, N_T)}$.

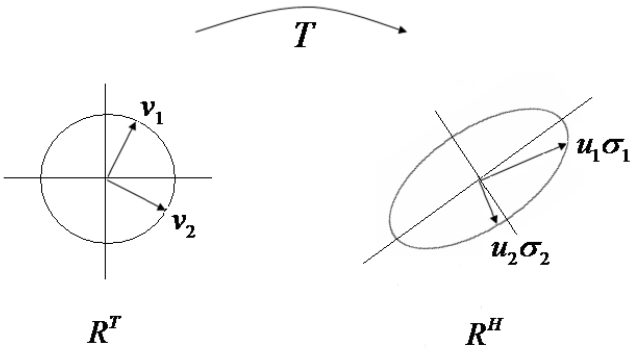


Figure 4.34: SVD concept illustration

The base vectors in U and V provide important directions in each subspace and the singular values σ_i provide a ranking of these directions. Looking at figure 4.34, it is clear that the result of T on the direction u_i is negligible if σ_i is so. The direction i can then be removed from the basis of both spaces (σ_i set to 0) with minor changes on the representation of T.

In other words, given a coupled representation of spaces $\mathcal{R}^T$ and $\mathcal{R}^H$ through T, the SVD decouples information of both spaces by generating new bases for each, which can be seen as the principal directions of T. Using a tolerance ϵ, it also gives the space of minimum size for which the altered application is close to the full one with an error ϵ on the operator. This minimum dimension is often called *effective rank*.

For the applications considered here, T is the response of a structure as function of time (transient response) or frequency (FRF). It is a basic fact of modal analysis, both numerical and experimental, that the effective rank of T is small.

Defining a $[Q]$ matrix storing a number of time steps as columns, it can be decomposed as

$$[Q] = [U]\,[\Sigma]\,[V]^T \tag{4.113}$$

With $[Q]$ of dimension $N \times Nt$, the SVD computes $[U]_{N\times N}$, $[\Sigma]_{N\times Nt}$ and $[V]_{Nt\times Nt}$. $[U]$ and $[V]$ are unitary matrices and $[\Sigma]$ is rectangular diagonal.

$[U]$ is a set of deformation shapes spanning the finite element space. Each shape u_j is associated to a value on the diagonal of Σ, σ_j characterizing its overall amplitude. $[V]$ spans the time space, as each v_j^T corresponds to the time variation of the associated u_j shape. The total response is thus given by

$$\{q(t)\} = \sum_j \{u_j\}\,(\sigma_j v_j^T(t)) \tag{4.114}$$

which is clearly very similar to the modal decomposition (4.103). The product $\sigma_j\,\{v_j\}^T$ is a vector amplitude very similar to the modal amplitudes $\{\alpha_j(t)\}$. Similarly a common result of modal analysis is that vectors $\{u_j\}$ are close to real or complex modes, which can be verified using correlation with the MAC or through the modal sensors introduced in section 4.4.3.

Visualizations of $[U]$, $[\Sigma]$ and $[V]^T$ on a limit cycle obtained for a simplified brake model are shown in figure 4.35.

Performing the SVD on the displacement time response yields Euclidean normalized shapes. Common problems involving the Euclidean norm in finite elements is a strong mesh dependence and the lack of link to common energies. More mechanically relevant shapes can however be obtained by applying the SVD on the strain or kinetic energy response [45].

In such case, and taking the strain energy as an example, the SVD writes

$$[Q]^T\,[K]\,[Q] = [U_K]\,[\Sigma_K]\,[V_K]^T \tag{4.115}$$

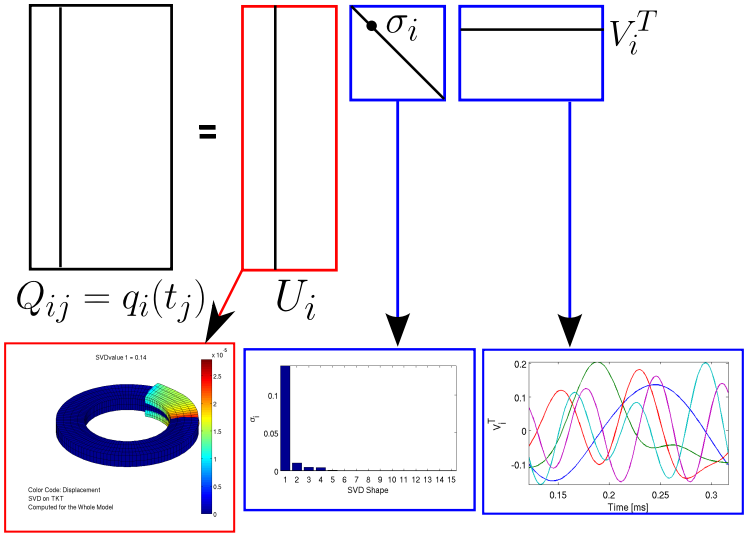


Figure 4.35: Illustration of the space-time decomposition concept using an SVD of the time response

The matrix size on which the SVD is performed is then only time step dependent, which in considered applications will be much smaller than the displacement response. $[U_K]$ is the basis of eigenvectors of the symmetric semi-definite matrix $Q^T K Q$ and shapes in physical space are recovered using

$$[U_{Strain}] = [Q]\,[U_K] \tag{4.116}$$

4.5 Application to the industrial model

This section presents a series of applications of the tools presented in the previous section. The focus is placed on providing methodologies that could be practical to industrial applications.

Modal damping is first validated in section 4.5.1. It is here exploited to check other potentially unstable modes in section 4.5.2, and to illustrate a suggestion of time simulation initialization strategy, in section 4.5.3. Finally, the limit cycles are decomposed by the SVD and correlated to the system real modes in section 4.5.4.

4.5.1 Modal damping, illustration and validation

The use of modal damping can be a powerful tool in design studies as it can be used to assess the effect of enhancing dissipation for specific modes. In chapter 5 the concept will be extended to the assessment of enhancing damping of components rather than system modes.

To illustrate the effect of modal damping and validate its implementation, three sample test cases are proposed, targeting the main unstable mode of the system, mode C44 (cf. section 3.4.2) through its real mode contributors. Figure 4.36 presents the effect of increasing the damping of mode R45 from 0 to 10%. Although the limit cycle does not seem to be reduced (cf. chapter 5), the growth rate is clearly reduced.

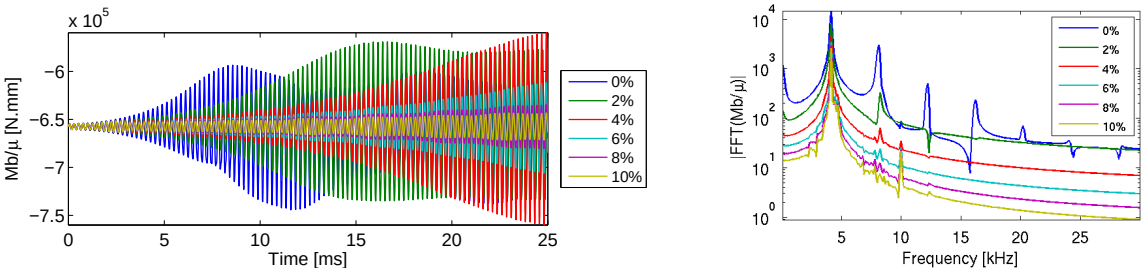


Figure 4.36: Braking torque and its FFT between 1 and 10ms as function of mode R45 modal damping for an initial excitation based on the original unstable mode C44

Modal sensor results are directly showing the response of the shapes impacted by the modal damping strategy. Figure 4.37 shows in particular the effect on modes R44 and R45. Using a logarithmic scale, the growth rates can be identified as the initial slopes observed. The higher the damping ratio, the lower the growth rate. The duality between modal damping and modal sensors is thus illustrated.

It is interesting to note that although mode R45 is the only one damped, the growth rate of mode R44 is also affected. In fact in the first 2ms of the simulation, mode R45 even shows decay. Its increase afterwards is then due to a more complex mechanism. This illustrates the fact that the instability couples both these modes, which will be extensively studied in chapter 5.

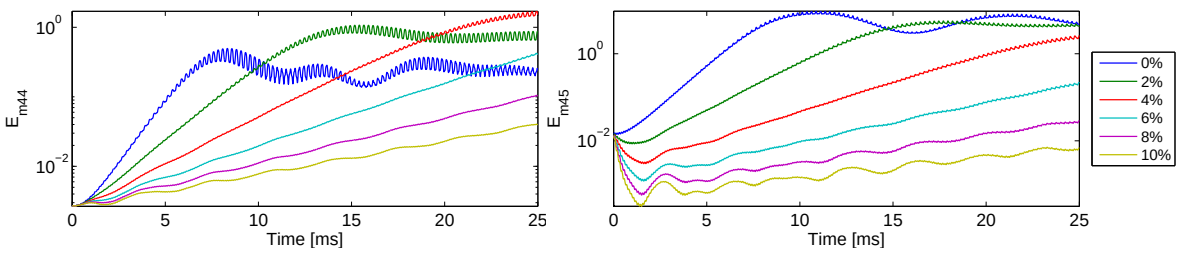


Figure 4.37: Modal sensor for modes R44 and R45 as function of mode R45 damping

One now seeks to asses the possibility to kill instability through modal damping. If all real mode contributors to a given unstable complex mode are damped over a threshold ratio, the complex mode should not react anymore. This is tested in the case of mode C44 shown in figure 4.38. The table giving the real mode amplitudes clearly indicates visible contributions from modes R42 to R48 only.

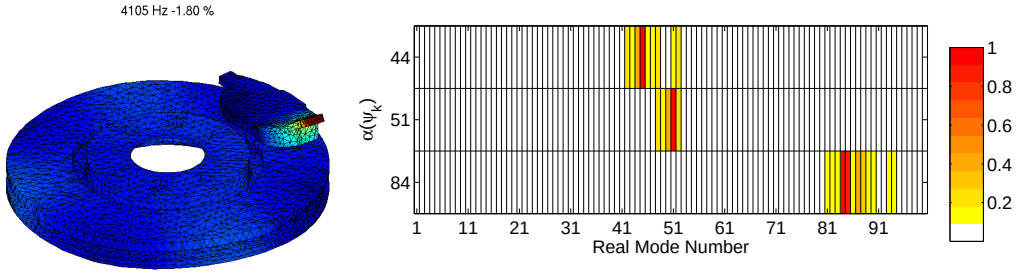


Figure 4.38: Unstable mode C44 and its projection on the real modes

A modal damping ratio of 10% is now applied, so that the 10% pressure step increment does not trigger the instability previously observed in the time simulation span, as plotted in figure 4.39.

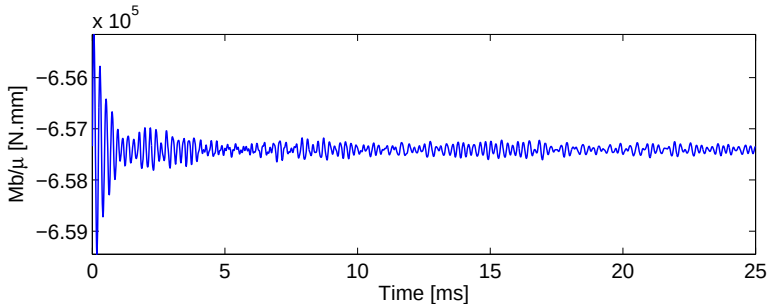


Figure 4.39: Braking torque for a 10% modal damping of modes R42 to R48

The contributors of mode C44 are assessed to be between modes R42 to R48. Their corresponding modal sensors are plotted in figure 4.40. From the initial excitation given by the pressure step, a decay rate is observed in the first 2ms. Modal damping has then an influence on the propensity of an unstable complex mode to grow in the response.

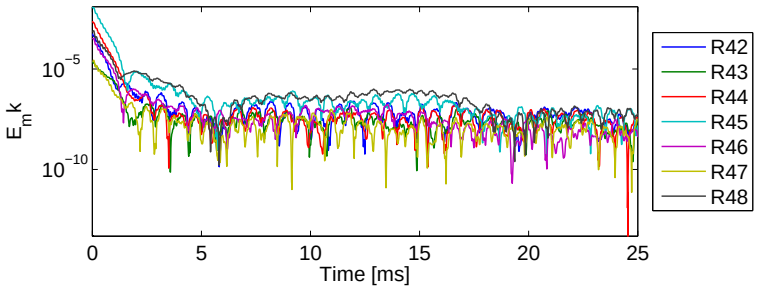


Figure 4.40: Modal sensor R42 to R48 for a 10% modal damping of modes R42 to R48, corresponding to the time simulation of figure 4.39

4.5.2 Modal competition, assessment of other unstable modes

As the system features several unstable modes, a so-called *modal competition* occurs in which unstable modes emerge as unstable, while others do not show unstable behavior in the simulation time scope. At 12 Bar, three main modes were detected as unstable, C44, C51 and C84 out of which only mode C44 appears when the structure is excited by a pressure step.

Using a complex mode target strategy allows to observe the behavior of other unstable complex modes. It is here performed by applying an initial velocity in the direction of a given complex mode shape. Figure 4.41 and 4.43 plot time simulation results initiated by an initial velocity respectively from mode C51 and C84 trajectory.

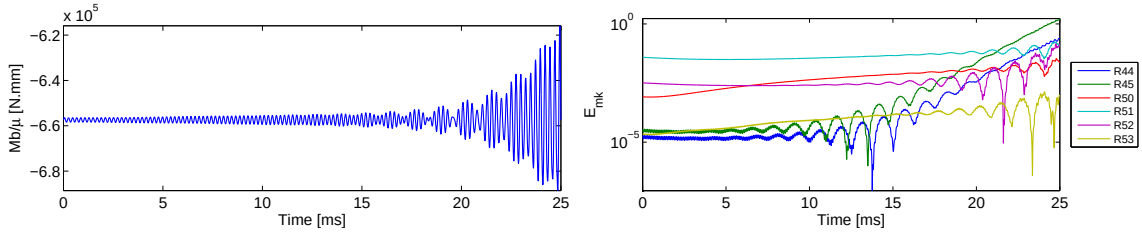


Figure 4.41: Targeting unstable mode C51, braking torque and main modal mechanical energies

Despite a favorable initial energy level, unstable mode C51 does not show an energy increase from its main contributors in figure 4.41. Mode C44 represented by modes R44 and R45 takes over the divergence phase after 15ms. This can also be observed in figure 4.42 were although the frequency linked to mode C51 is visible, the frequency linked to C44 and one harmonic appears in the second part of the simulation.

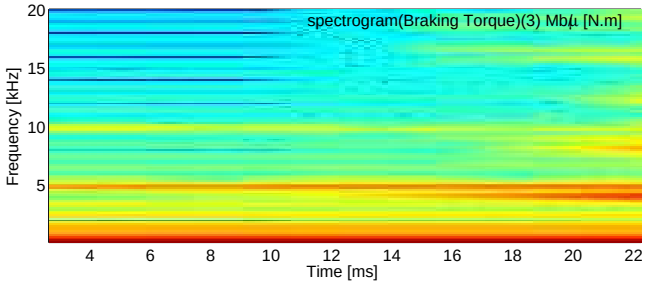


Figure 4.42: Targeting unstable mode C51, braking torque spectrogram

The energy levels between mode C84 and mode C44 are more different than between mode C51 and mode C44. Figure 4.43 shows a global decrease from the initial levels. Again, mode C44 takes over, which can be seen in the modal mechanical energies, figure 4.43 and the spectrogram of figure 4.44.

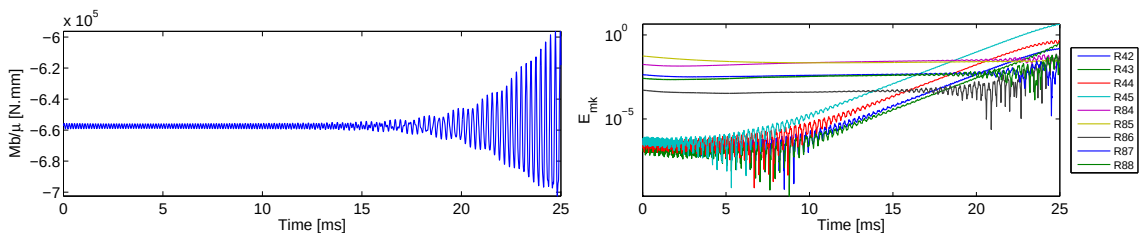


Figure 4.43: Targeting unstable mode C84, braking torque and main modal mechanical energies

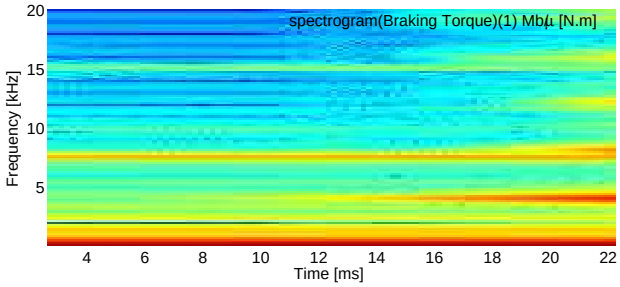


Figure 4.44: Targeting unstable mode C84, braking torque spectrogram

From these observations, mode C44 wins the modal competition. The question is however to decide whether modes C51 and C84 are really unstable or not. A test is performed where real modes R42 to R48 contributing to mode C44 are damped at 10% so that the main instability disappears. A new simulation with a neutral initialization through a 10% pressure step on the first two time steps is then launched.

Figure 4.45 shows that the initial excitation first decreases, which is coherent with the modal mechanical energy results. Indeed, the modes most excited by the pressure step seem at a rather high frequency and look stable. A slight increase of the braking torque is spotted in the end of the simulation (much longer than in section 4.5.1) which is confirmed by the modal mechanical energies, where modes R83 and R84 show a clear exponential increase.

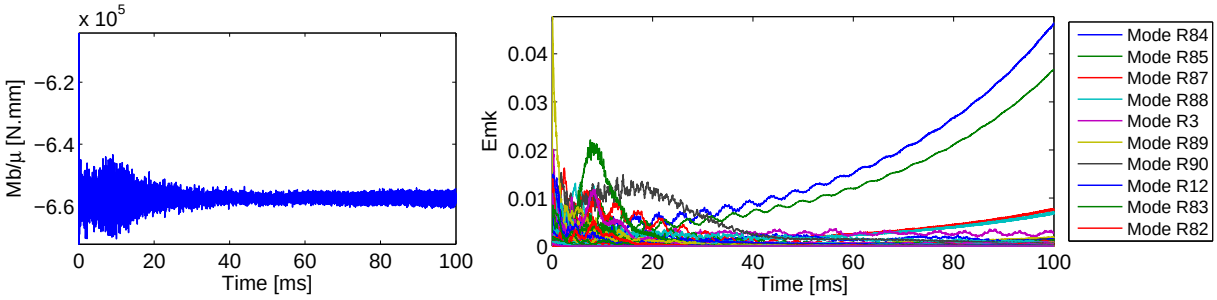


Figure 4.45: Brake squeal simulation at 12 Bar, braking torque and main modal mechanical energies. Modes R42 to R48 damped at 10%, initialization by a 10% pressure step

Although mechanical energies and vibration levels are low, the instability exists. The displacements sensors on the disc surface plotted in figure 4.46 ranges in tenth of microns, which is indeed low. A follow up to this simulation could be needed to estimate final vibration levels.

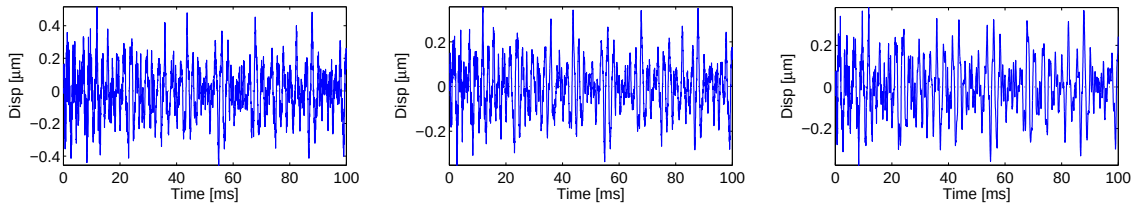


Figure 4.46: Brake squeal simulation at 12 Bar. Displacements obtained on the disc ODS at 3 sample points. Mode C44 damped at 10%, initialization by a 10% pressure step

The presented results show that all detected instabilities in the frequency domain may potentially be unstable in the time domain. By damping the first mode winning the *modal competition*, another mode shows up as unstable. Time simulation is therefore a good way to verify which mode wins the modal competition and whether other modes appear later in the simulation once energy levels are higher.

4.5.3 Initialization strategy

The growth phase of time simulations tends to represent a major part of the simulation cost. At the end of the simulation from figure 4.24 in section 4.3.2, mode R45 clearly becomes preponderant although the initial excitation was neutral (pressure step).

In a design phase where limit cycles already identified need to be reanalyzed, the objective of this section is to propose and analyze initialization strategies that limit the numerical cost of the long transition towards the bounded part of the response. The idea tested is to set an initial velocity in the direction of a given mode shape to influence a particular energy distribution from the beginning of the simulation, as already used in section 4.5.2. This initialization methodology is clearly not physical, but the *numerical* trick is very useful to analyze multiple configurations.

Table 4.1 presents the amplitudes obtained from modeshape C44 through the use of a multiplicative factor taken as a comparison reference in the following.

Amplitude Factor $[s^{-1}]$	Max. Vel. $[mm.s^{-1}]$
25	1.1742
50	2.3483
100	4.6967
200	9.3933
500	23.483
700	32.878

Table 4.1: Initial velocity amplitudes to trigger mode C44

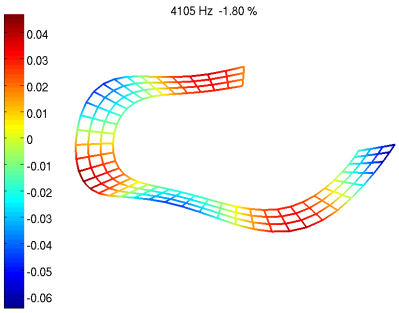


Figure 4.47: Mode C44 on ODS

Figure 4.48 shows that a single limit cycle is obtained for all excitation amplitudes while just reducing the transition time until a threshold value. Above this threshold, the system disturbance is too important, which causes a longer transient response or even divergence.

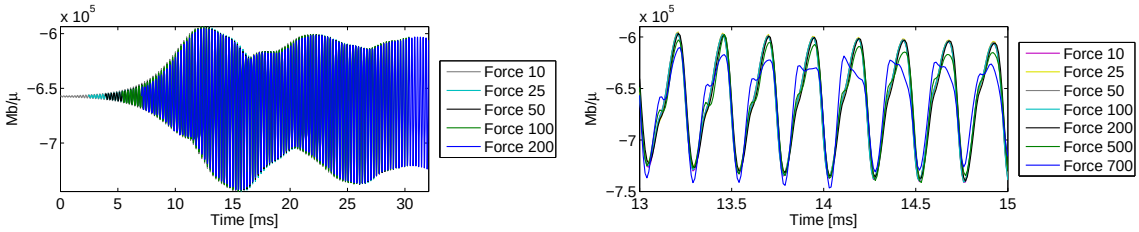


Figure 4.48: Real brake squeal simulation as function of the mode C44 excitation amplitude. Signals are shifted in time according to the maximum level instant

Modal sensors lead to the same conclusions in figure 4.49, but allow finer observations on the initial velocity amplitude. Indeed, there is a proportional relationship between the initial velocity and the mechanical energy. The initial displacement is null so that

$$E_{mk0} = \frac{1}{2}\dot{\alpha}_{k0}^2 = \frac{1}{2}A_f^2 \Re\left(\beta_k(\psi_l)\right)^2 \tag{4.117}$$

where A_f is the amplitude factor and β_k the projection of targeted complex mode ψ_l on real mode ϕ_k.

For energy levels over a threshold (seemingly the established limit cycle amplitude as seen in figure 4.48) the divergence phase does not exist and is replaced by an initial decrease and transient oscillations around the final value.

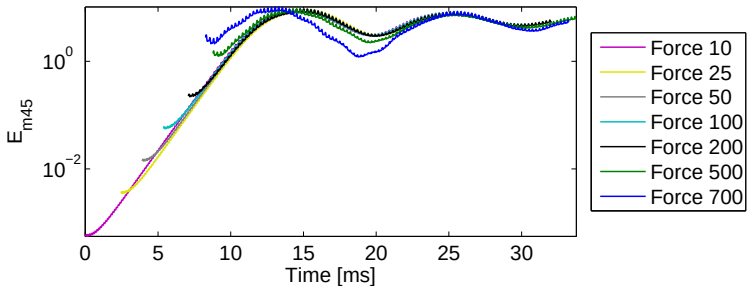


Figure 4.49: Mode R45 mechanical energy contribution to the response as function of the initial modeshape C44 velocity amplitude. Signals are centered according to the maximum level instant.

As a perspective, *a priori* estimations of the limit cycles could be obtained by the non linear analyses performed in section 3.4.2. The computational cost advantage is however arguable. Continuation is another strategy to limit computational costs and this will be detailed in chapter 5.

4.5.4 Limit cycles and SVD

The installation of limit cycles is obtained from the initial simulation of figure 4.24. A 75ms continuation was then performed to obtain a complete 100ms simulation, whose result is presented in figure 4.50. After a modulation period, the signal becomes very stable.

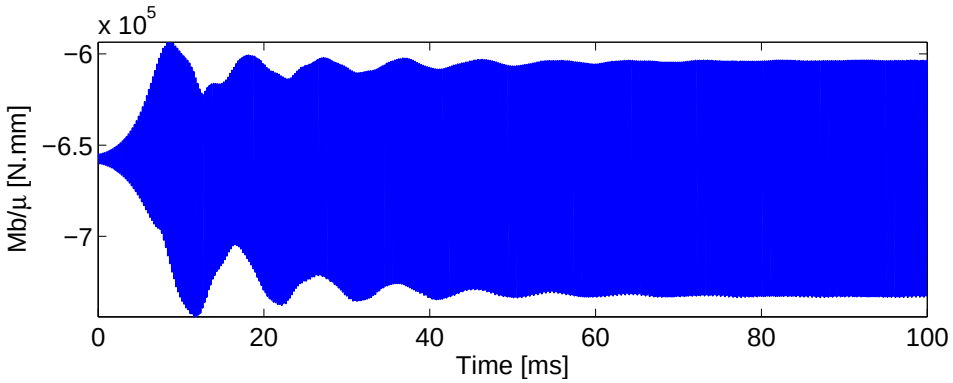


Figure 4.50: 100ms of the real brake squeal simulation at 12 Bar, initialized by excitation of mode C44

A slip-separation pattern occurs. Gap plots presented in figure 4.51 show two main evolution trends. Either separation patterns occur or the oscillation is well under zero gap. Due to the exponential contact law, the gap evolutions are very smooth.

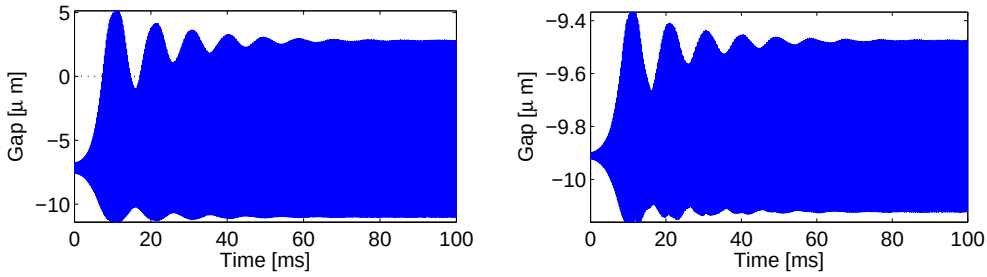


Figure 4.51: Gap computed at sample contact points, one showing a slip-separation pattern (left) and one far from separation (right)

As shown in the spectrogram figure 4.52, the harmonics of the main instability frequency appear in the simulation. This highlights the non linear effects in the simulation.

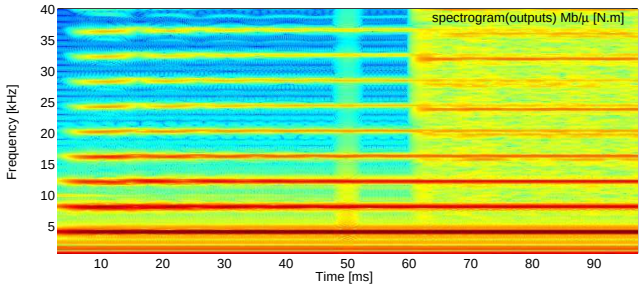


Figure 4.52: Spectrogram of the brake squeal simulation at 12 Bar over 100ms

The non-linear character of the limit cycles is illustrated by the modal sensors results shown in figure 4.53. One notes for example the apparition of mode R95 in the response. It is a stable mode at $8kHz$ excited by the second harmonic of the 4kHz response.

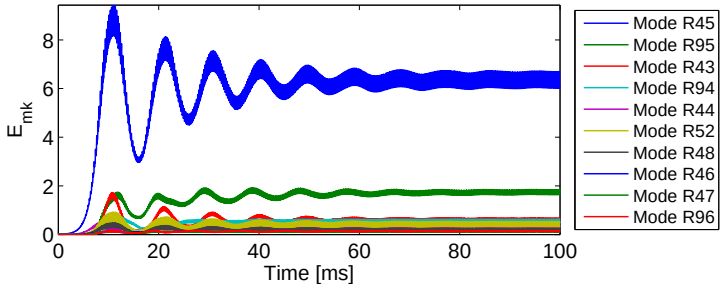


Figure 4.53: Modal sensors results for the 100ms simulation at 12 Bar, initialized by excitation of mode C44

To represent the total mechanical energy of the system, a lot of real modes are needed, mode R45 only contributes by 55% of the total mechanical energy. This motivates the use of an analysis tool independent from the system modes.

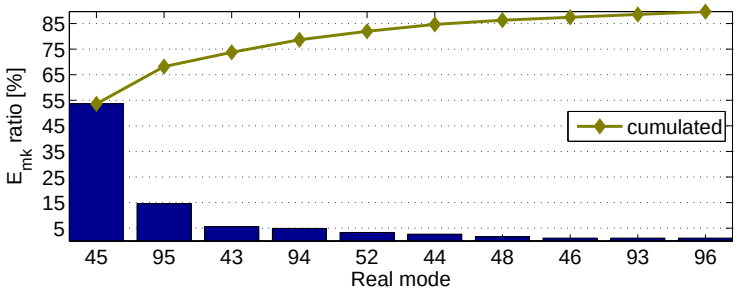


Figure 4.54: Modal mechanical energy ranking by ratios to the total one

Non-linear aspects can also be observed on phase plots, which are commonly performed in the literature of non-linear analysis. Sample points are taken over the surface contact of the outer pad for this purpose. The selected points are shown in figure 4.55.

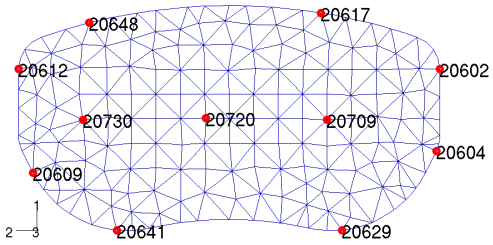
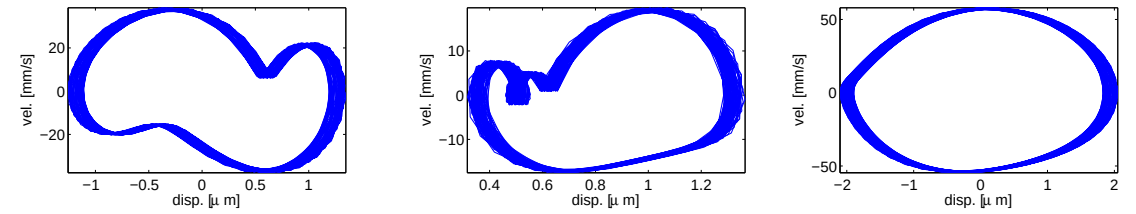


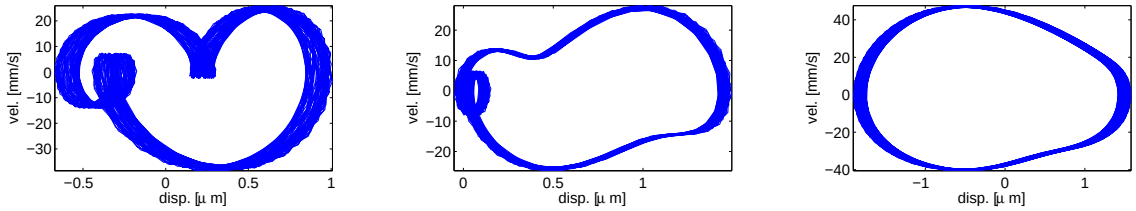
Figure 4.55: Selected pad nodes for the following phase plots. The disc is moving from left to right. Contact normal direction 3, contact plane, directions 1 and 2

The phase plots are taken between 80ms and 100ms as it corresponds to the moment the limit cycle is stabilized. The front side points 20,602 and 20,604, figure 4.57m-r show large displacements and velocities in comparison to the rear and middle sides. Their cycles are rather simple and close to a line. The rear and middle side points show lower velocities but more complex movements, in the contact plane for the rear points, figure 4.56a-o and in the normal direction for the middle points, figure 4.57a-l.

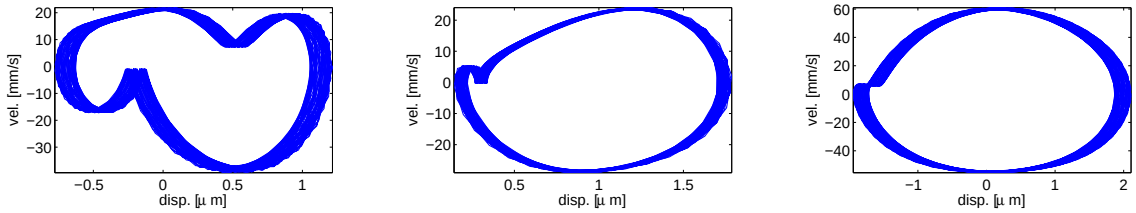
The multi-harmonic response observed here can be compared to the work of Coudeyras *et al.* [115]. In such case, the pseudo-periodic response can be characterized by the limit cycle envelopes observed. The degree of complexity of the response however cannot be assessed by such post-treatment.



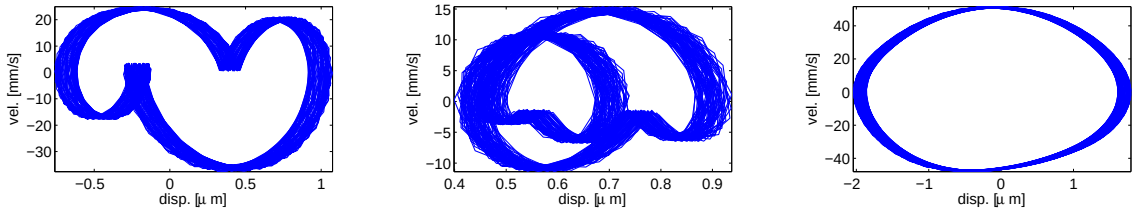
(a) Point 20612, dir 1 (b) Point 20612, dir 2 (c) Point 20612, dir 3



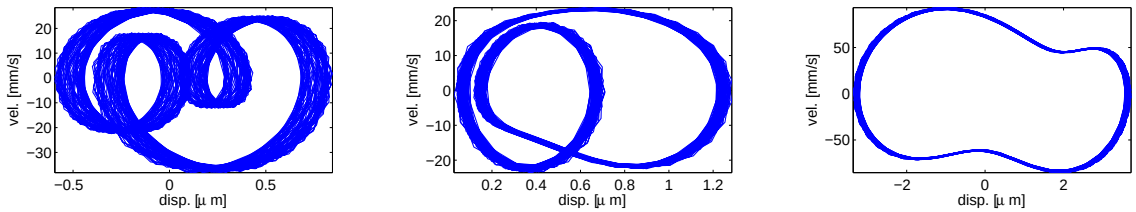
(d) Point 20609, dir 1 (e) Point 20609, dir 2 (f) Point 20609, dir 3



(g) Point 20648, dir 1 (h) Point 20648, dir 2 (i) Point 20648, dir 3

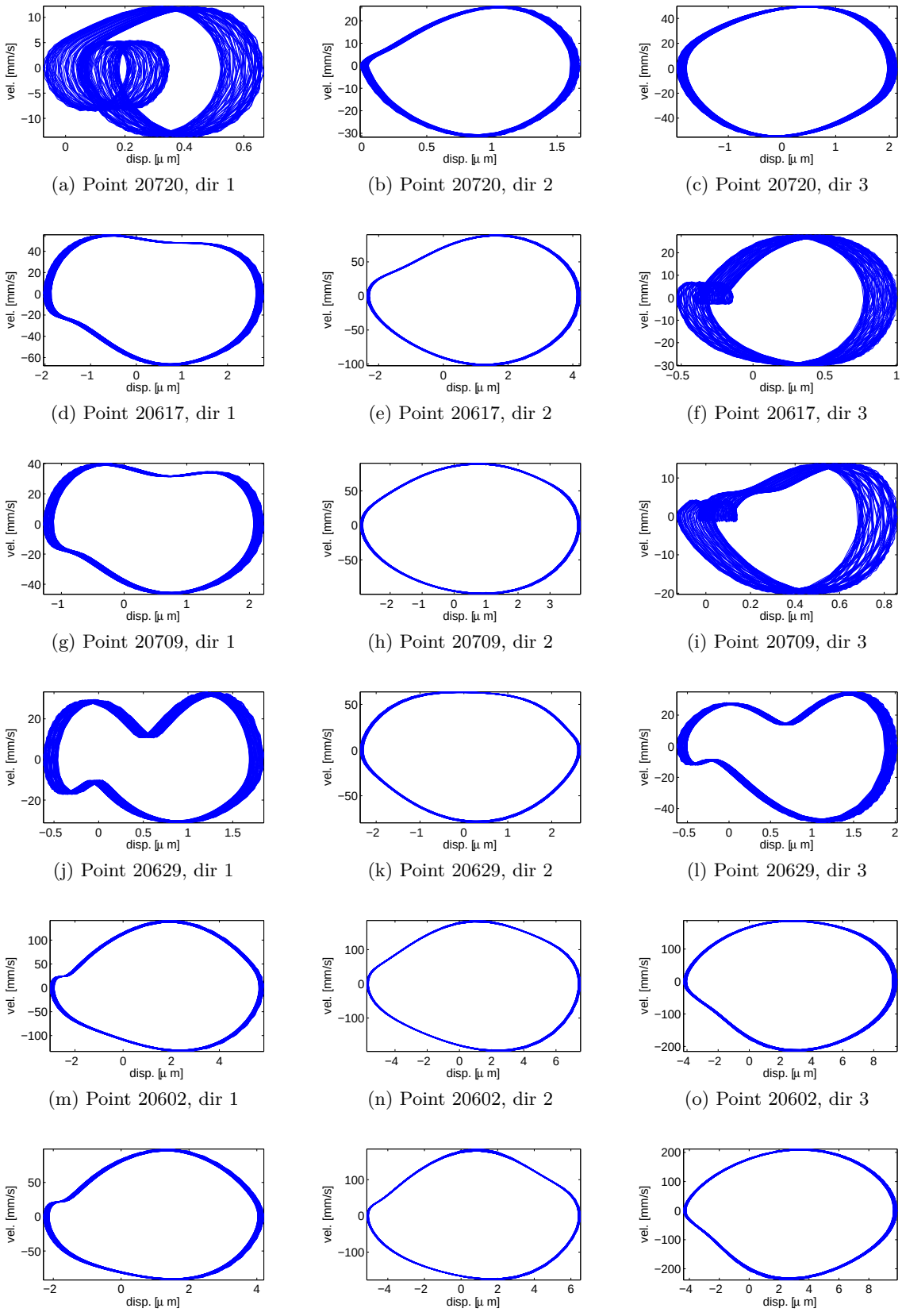


(j) Point 20730, dir 1 (k) Point 20730, dir 2 (l) Point 20730, dir 3



(m) Point 20641, dir 1 (n) Point 20641, dir 2 (o) Point 20641, dir 3

Figure 4.56: Phase plots of sample points from the outer pad surface (rear side), between 80 and 100ms of the simulation of figure 4.50



(a) Point 20720, dir 1 (b) Point 20720, dir 2 (c) Point 20720, dir 3

(d) Point 20617, dir 1 (e) Point 20617, dir 2 (f) Point 20617, dir 3

(g) Point 20709, dir 1 (h) Point 20709, dir 2 (i) Point 20709, dir 3

(j) Point 20629, dir 1 (k) Point 20629, dir 2 (l) Point 20629, dir 3

(m) Point 20602, dir 1 (n) Point 20602, dir 2 (o) Point 20602, dir 3

(p) Point 20604, dir 1 (q) Point 20604, dir 2 (r) Point 20604, dir 3

Figure 4.57: Phase plots of sample points from the outer pad surface (middle to front side), between 80 and 100ms of the simulation of figure 4.50

Unlike modal sensors, the SVD allows to find the displacement principal shapes, regardless of the modal characterization. Although over 10 real modes are necessary to represent the system displacement, as shown in figure 4.54, the space spanned by the limit cycle is only of dimension two. Figure 4.58 indeed plots the amplitude and time variation over the limit cycle of the principal shapes computed following the formulation provided in section 4.4.4. 99% of the displacement is here described by the combination of two shapes, plotted in figure 4.59.

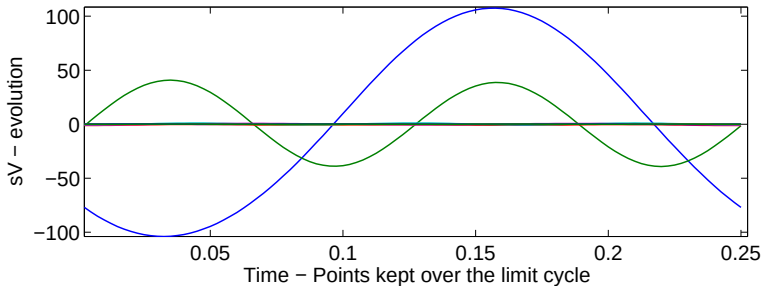


Figure 4.58: Time variation and amplitude of the shapes given by the SVD in the limit cycle

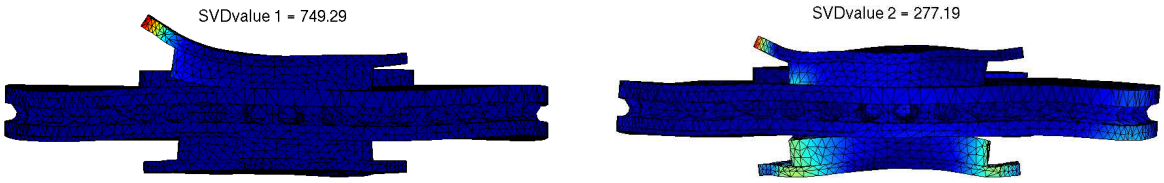


Figure 4.59: First two shapes given by the SVD on the squeal limit cycle

The first shape is close to mode C44 as the MAC shows a correlation of 74% with it in figure 4.60. The second shape shows large displacements of both pads and the disc, it is correlated at 65% to mode C95 by the MAC. The MAC results are also in good coherence with the modal sensors results of figure 4.53. Mode R45 and R95 appear as predominant, along with others modes around these main peaks.

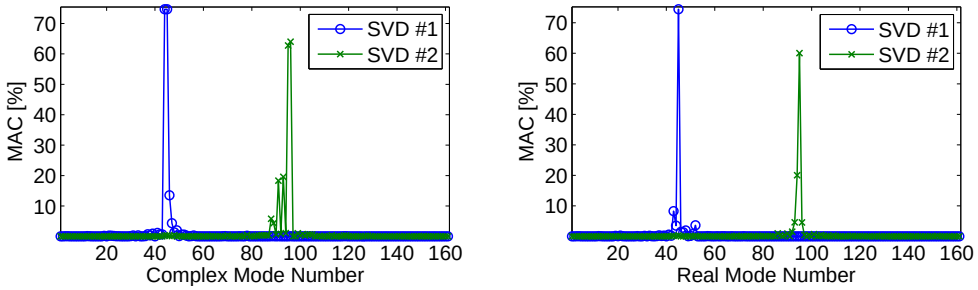


Figure 4.60: Correlation between the principal shapes given by the SVD and the brake steady state complex (left) and real (right) modes through a MAC computation

The principal shapes provided by the SVD can therefore be used to characterize the system in the limit cycle. Their evolution through time is given in figure 4.58. Their evolution is almost periodic, the first shape signal is at the main instability frequency of $4kHz$, while the second one is at a frequency of $8kHz$. This is also coherent with the MAC results of figure 4.60. The integrated exploitation of the tools illustrated in this section is eventually presented in chapter 5.

5

Design studies for the industrial brake

Contents

5.1 Limit cycles, towards a better understanding

The time simulation post-treatment tools presented in section 4.4 are here applied to study in depth the limit cycle obtained in the real brake time simulations of section 4.3. The objective is to gain an understanding of the mechanics of the observed instability, which produced complex cycles. Modal interactions and energy transfers in the limit cycle are highlighted.

5.1.1 In depth analysis of the limit cycle

The squeal limit cycle obtained in the simulation presented shows mechanical energy variations in the cycle. In particular, figure 5.1 shows that the modal mechanical energy of mode R45 varies by 11% in amplitude in the cycles. Besides, unstable mode C44 represented by real modes R43, R44 and R45 only shows 65% of the total modal mechanical energy, as shown in figure 4.54. One will thus seek to analyze the importance of non modal energy.

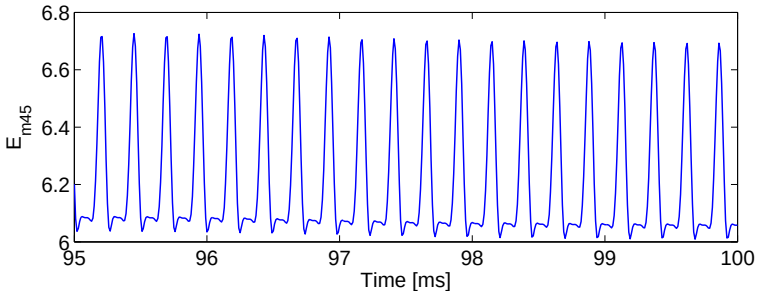


Figure 5.1: Mode R45 sensor at the end of the squeal simulation

Modal sensors have been introduced in section 4.4.3 by writing the displacement as the sum of the nominal modes contribution q_ϕ and an orthogonal part q_r, such that

$$q(t) = q_\phi(t) + q_R(t) \tag{5.1}$$

The orthogonal part represents a deviation from the initial state, whose mechanical energy contribution can be computed by equation (4.109).

The deviation from the steady state real mode space is limited. Figure 5.2 shows in particular that the total mechanical energy is close to the sum of the modal mechanical energy contribution of the first 250 modes, as the difference is on average around 0.7% of the modal mechanical energy.

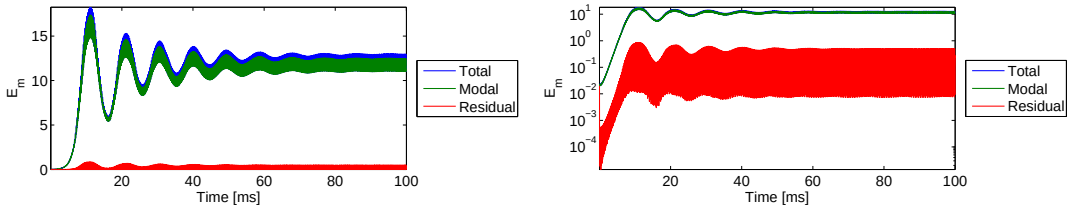


Figure 5.2: Total mechanical energy from the first 250 real modes against the total mechanical energy

The mechanical energy not contributed by the real modes varies through time from almost nothing to 4% of the total. Figure 5.3 shows in particular that the highest differences are located at the high energy peaks. Such variation shows that the dynamic expression of the system actually evolves significantly as function of the time.

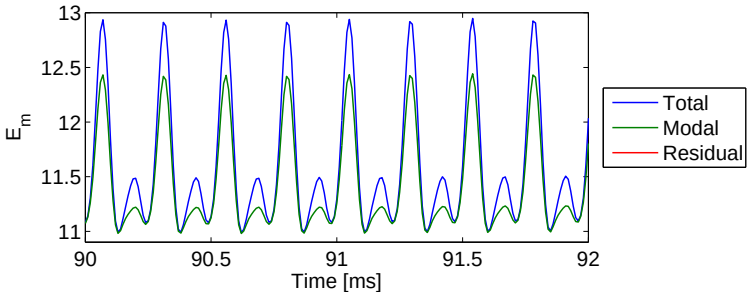


Figure 5.3: Differences between the modal mechanical energy and the total mechanical energy in the squeal limit cycle

The definition of the total mechanical energy in figures 5.2 and 5.3 is not conventional, as it was based on the mechanical energy of the system tangent to the non-linear one at the time origin. This explains the high peaks observed.

A more conventional mechanical energy computation is provided in figure 5.4. The energy due to the contact penalization is there integrated exactly at each time step, noting that the contact force at each contact point x_i is

$$f_N(x_i) = \frac{\partial E_i}{\partial g_i} = \frac{\partial}{\partial g_i}\left(\omega_i J(x_i)\frac{p_0}{\lambda}e^{-\lambda g_i}\right) \tag{5.2}$$

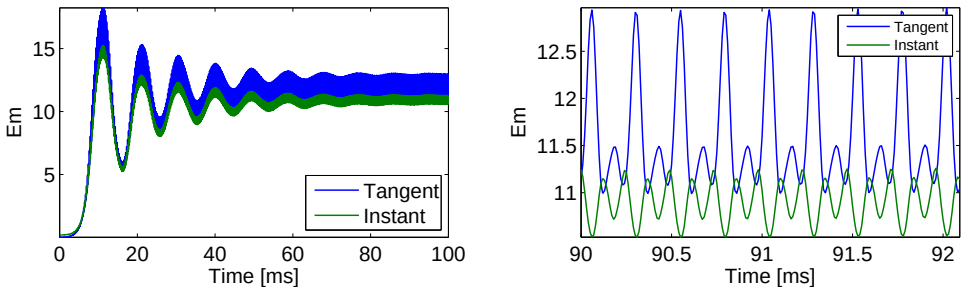


Figure 5.4: Actual mechanical energy of the non-linear system (*Instant*) against the total mechanical energy of the tangent system (*Tangent*)

The zoom in on the simulation end in figure 5.4 right shows that the tangent model is not perfectly representative of the transient states in the limit cycle. The differences are explained by the fact that in the case of separation, high strain energy is seen by the tangent system for which the contact is a linear spring, although the separation does not generate any contact strain in reality. In the case of higher compression, the tangent state is underestimating the strain energy as the rigidification is not taken into account.

It can be seen nevertheless that the exact total mechanical energy is oscillating over time. The system thus shows a sequence of energy dissipation phases and energy generation phases. The light will be cast on this pattern later on.

Since the initial tangent state is not representative of the system in the limit cycle, it can be guessed that the system modal properties evolve through time due to the contact non linearity. Using analytical expressions such as the exponential formulation for the contact laws allows the computation of transient tangent contact states with their corresponding contact stiffness, as presented in section 3.4.2 for pseudo-cycles.

In application to the limit cycle presented in this section, the contact stiffness evolution is plotted for both pads in figures 5.5 and 5.6. The contact stiffness is computed at the Gauss contact points triangularized by a 2D Delaunay method for visualization purposes, for a deformation sequence taken in the limit cycle.

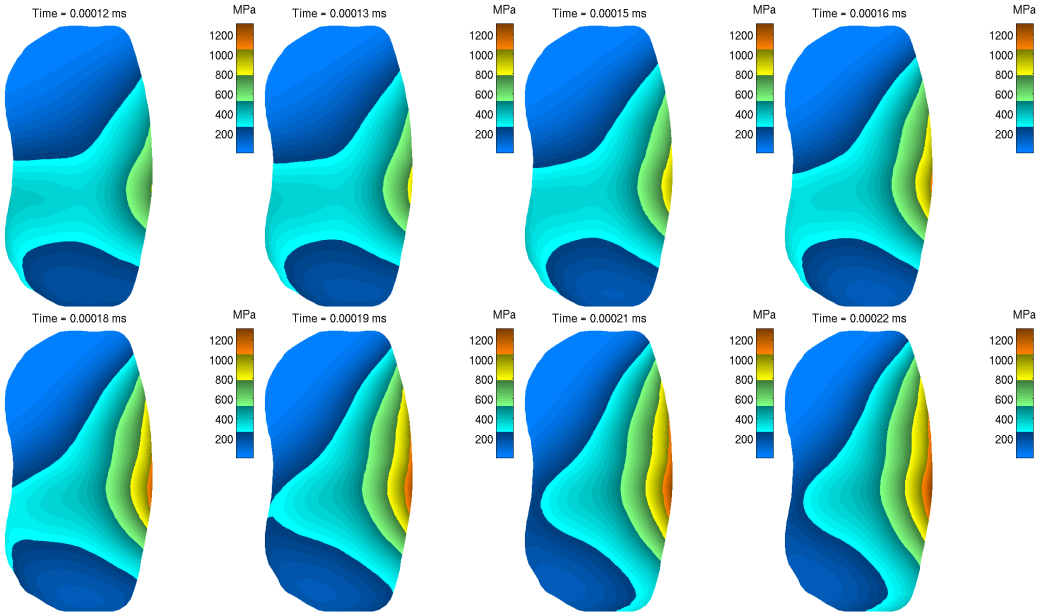


Figure 5.5: Inner pad contact stiffness cartography variation in the squeal limit cycle

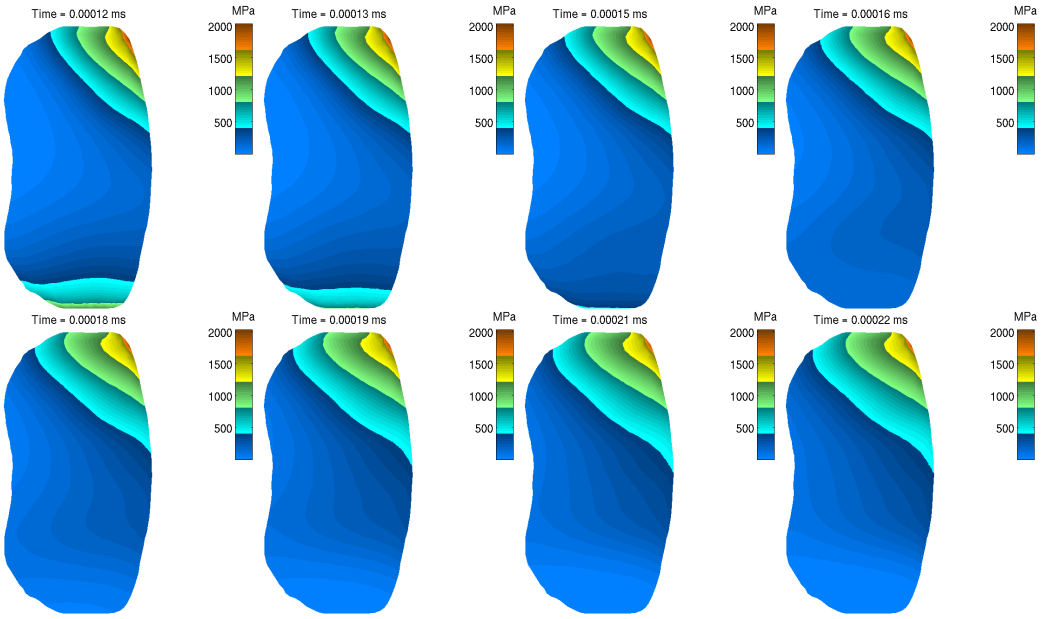


Figure 5.6: Outer pad contact stiffness cartography variation in the squeal limit cycle

Great differences exist between both pads due to the floating piston implementation. The inner pad, figure 5.5, is directly moved by the piston which is located at the pad middle close to the outer side. The inner pad is thus mostly loaded at its center with both pad ends only lightly loaded. The outer pad, figure 5.6, is moved by the caliper fingers which loads the pad ends and unloads its center. Due to the disc rotation, the outer pad entry is more loaded than the exit in a classical arching shape. These different load applications yield large contact stiffness differences up to 60% between both pads.

During the squeal limit cycle, large contact stiffness variations occur. The inner pad shows a 66% variation of the maximum contact stiffness under the piston while the load is more or less distributed on the pad. The disc entry side is not very loaded.

The disc entry side of the outer pad does not vary much and keeps highly loaded. The disc exit side however shows large variations as a slip-separation pattern occurs.

The important stiffness variation at the pad/disc interface thus changes the system modal properties and explains the variations observed in the modal mechanical energy plots of figure 5.3. The fact that the pad/disc interface is not reduced is deemed sufficient to limit errors.

Having shown that the tangent stiffness has significant variations, the usual assumption that the system tangent to the static state is most relevant should be questioned. Dynamic stability diagrams are thus realized. The system complex modes at several deformation states in the limit cycle are computed (like for the pseudo-cycles). It appears that mode coupling/coalescence patterns occur and some modes show great frequency variations. Figure 5.7 plots frequency variations over time at 4kHz (mode C44 frequency) and 5kHz (mode C51 frequency). The plot shows a mode with a frequency variation of 1kHz between 3.4 and 4.4kHz.

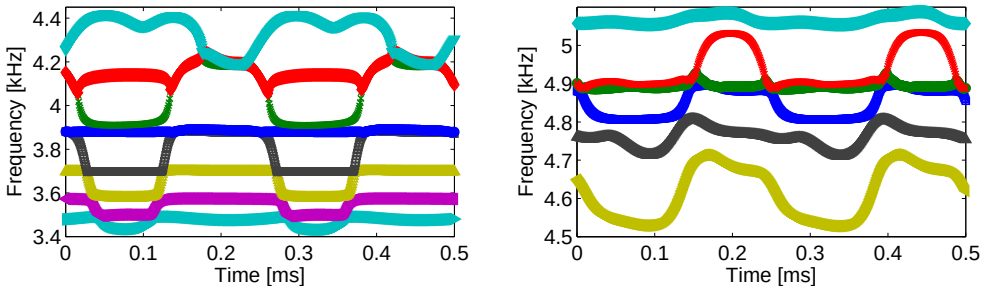


Figure 5.7: Frequency variation of the brake complex modes in the squeal limit cycle at 4 and 5 kHz

Complex mode damping also evolves through time and stable/unstable transitions occur. Figure 5.8 shows the dynamic stability diagrams of modes C44, C51 and C84, which are the most unstable. All three modes show such stability transitions. The system consequently passes through unstable states with energy increase and stable states with energy decrease. Figure 5.9 illustrates this pattern by showing the damping evolution of mode C44 along with modal frequency variations close to 4 kHz.

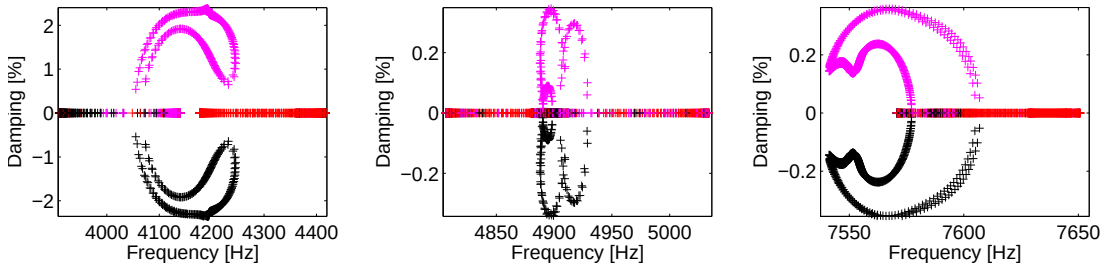


Figure 5.8: Dynamic stability diagrams for modes C44, C51 and C84 (left to right)

Mode C44 shows interactions with other modes and is really unstable only half of the time in the limit cycle. Figure 5.9 shows that after the first coalescence and until the last one, no instability occurs in the studied frequency band (3.2 to 4.5kHz), which is conform to the observations made in figure 5.8.

During the frequency/damping evolution, the shape itself is evolving. An attempt was made to track the shape along with the stability diagram computations. This is not trivial as using an automatic algorithm to detect frequency variations in such a large bandwidth is difficult. A significant number of points were however obtained, and results are plotted in figure 5.10. Shapes away from the pole veering points were easier to obtain.

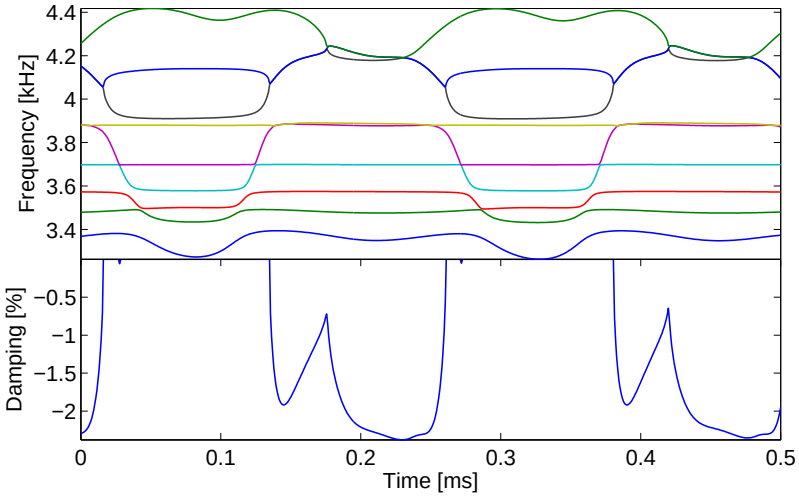


Figure 5.9: Evolution of mode C44 in the limit cycle. Frequency variation at 4kHz and corresponding damping evolution

Figure 5.10 shows the identified mode indices, and the MAC evolution between the identified shape and the nominal mode C44 shape. Obviously the MAC reaches 100% when the cycle passes through the initial deformation state. It can be seen that the shape evolves abruptly at coalescence points, and shows smooth evolutions in between. The shape is however not very well correlated to the initial complex modes during most of the limit cycle. An analysis of shapes is thus deemed useful.

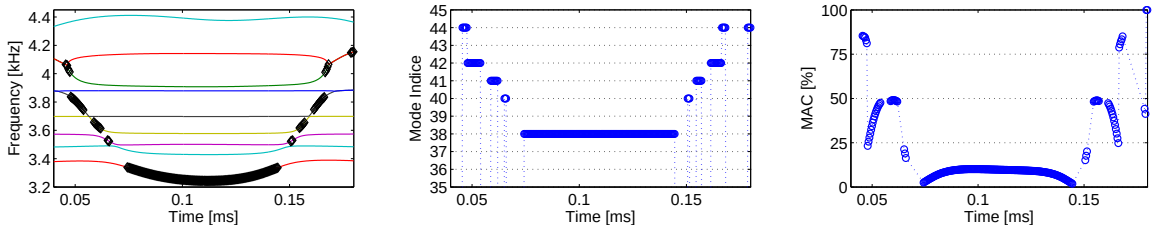


Figure 5.10: Identification of the moving shape in the frequency diagram, mode tracking indices and MAC against the nominal unstable mode

A subset of the identified points, visible in figure 5.11 is consequently chosen to observe the mode shapes.

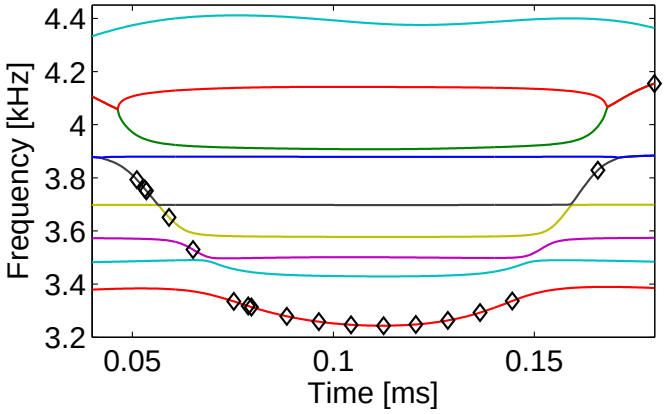


Figure 5.11: Selected points in the diagram for the shape deformation analysis

Although the MAC results are poor, the shapes shown in figure 5.12 mostly involves an outer pad rear bending motion at all times. The moving shape is thus visually correlated to mode C44. High motion amplitude of the anchor can be detected at the middle of the cycle (t around $0.1ms$) and is the source of the poor MAC results. This indicates that component contributions to the modeshape evolve during the cycle.

To analyze this evolution of component contribution, figure 5.13 shows for the identified shape, the component-wise strain energies, as function of time. The strain energies are ranked from highest in solid dark red to lowest in transparent light blue. The results clearly cast a new light on the disc/pad/caliper interaction.

At the nominal state, or as observed by SVD (section 4.5.4), most energy is in the disc and pad. The separation pattern responsible for the system stabilization (from 0 to $0.0752ms$) is a clear pad/disc interaction. The disc oscillation thus pushes the pad away. Once the pad starts separating, it constrains the caliper (from $0.0788ms$ to $0.12ms$), so that the interaction shifts towards a pad/caliper interaction. The caliper response pushes the pad back against the disc, which makes the system come back to its original unstable state.

The pad/disc/caliper interaction is thus sequential, and explains the shape differences. This analysis thus provides a complementary view to the stiffness cartography evolution plots of figures 5.5 and 5.6 and provides a physical explanation of the stabilization.

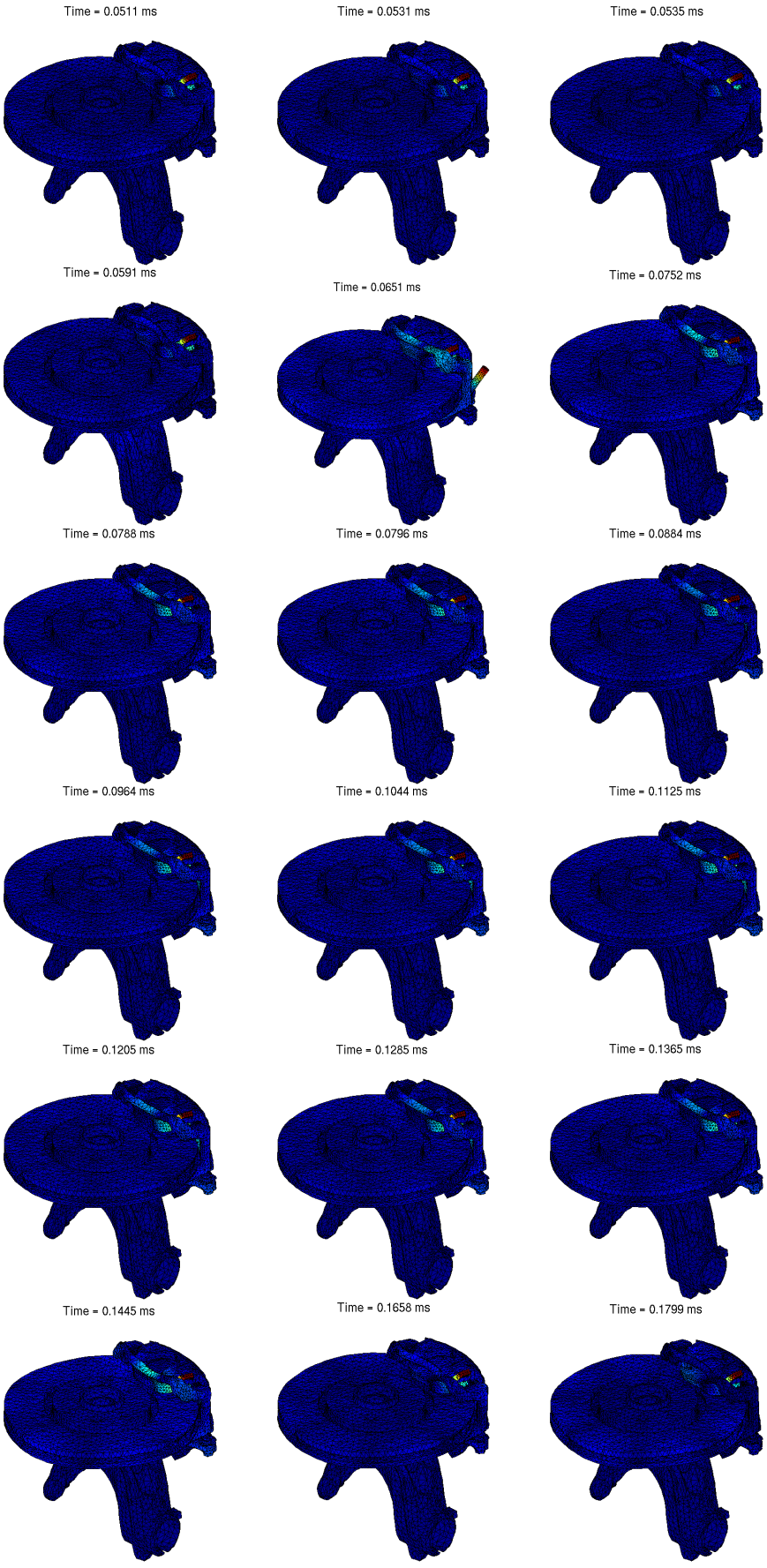


Figure 5.12: Identified moving shapes at several instants of the limit cycle, deformations plotted

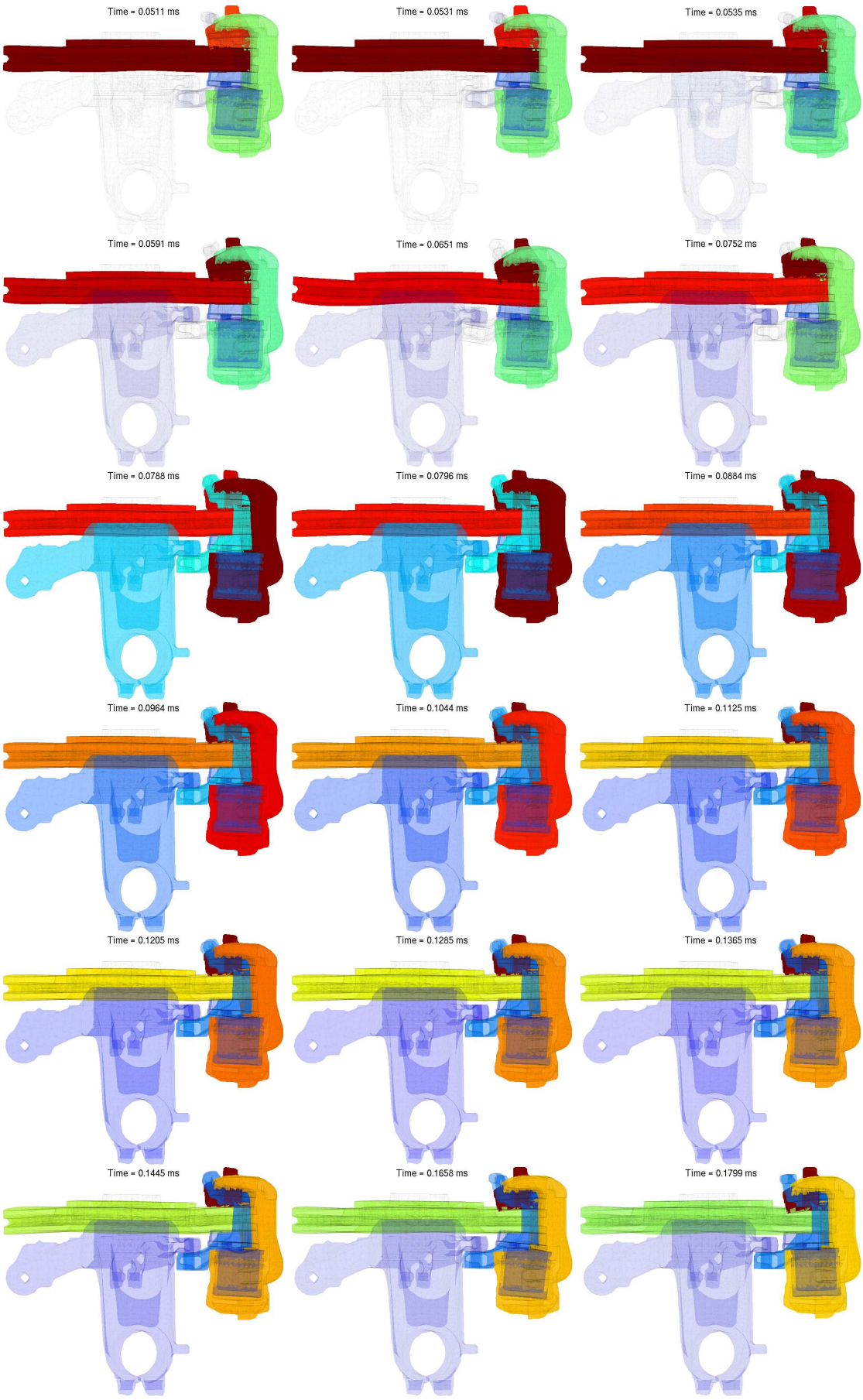


Figure 5.13: Identified moving shapes at several instants of the limit cycle, component-wise strain energy ranking. From high values in solid dark red to low values in transparent blue

5.1.2 Effect of damping on the instability mechanism

Damping treatment is a solution often considered to alleviate squeal problems. The ability to analyze damping effects is thus critical. The modal damping approach proposed in this work provides a simple mechanism for such studies. Evaluating modal damping effect from the initial steady state can however be tedious as a full limit cycle transient increase must be recomputed. As presented in section 4.5.1, the growth rate can decrease with modal damping, requiring potentially longer time simulations. The possibility studied in this section is to restart from a stabilized limit cycle. Illustrations are given and differences between restart analysis and full recomputation are analyzed.

To demonstrate the ability to use restarts, mode R44 damping at 2% is chosen as a case study. Results obtained over a 50ms simulation either from a restart or from the steady state are presented in figure 5.14. The limit cycles obtained match very well so that it is possible to overlay both braking torque signals.

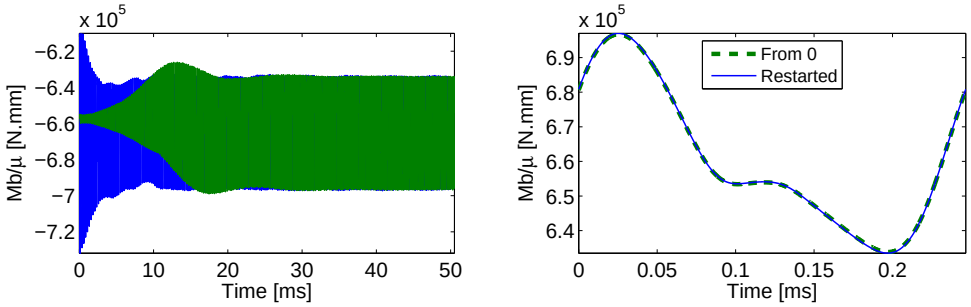


Figure 5.14: Difference evaluation between a limit cycle obtained from a restarted analysis with a 2% mode R44 damping or restarted from time 0. Left: new limit cycle establishment. Right: limit cycle identification

Similarly, figure 5.15 shows for modes R44 and R45 that energies almost overlay for both initial conditions. Finally figure 5.16 shows the cycles similarity by plotting the MAC between the principal shapes extracted from the SVD of both cycles. The limit cycles are identified at the end of both simulations, using a tolerance maximum amplitude trigger. The results are given in figure 5.14right. The MAC matrix shows an over 99% diagonal which really demonstrates the correlation between the final limit cycles.

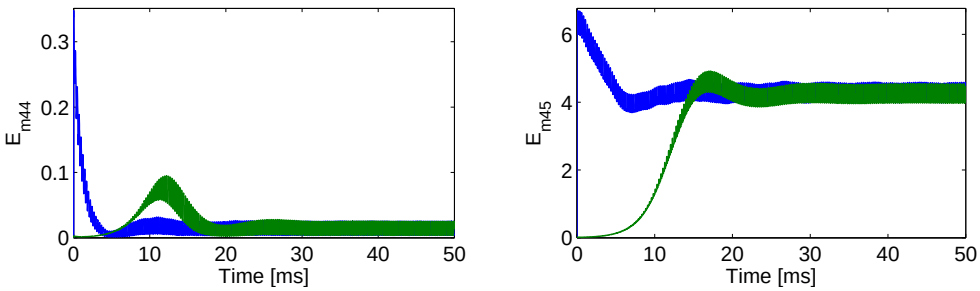


Figure 5.15: Difference evaluation between a limit cycle obtained from a restarted analysis with a 2% mode R44 damping or restarted from time 0. Modal sensor R44 and R45 results

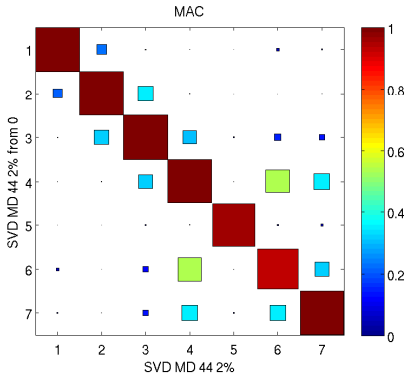


Figure 5.16: Difference evaluation between a limit cycle obtained from a restarted analysis with a 2% mode R44 damping or restarted from time 0 MAC between the shapes obtained by the SVD on the limit cycles

Based on these observations, the use of restarts is taken to be appropriate and will be used for further studies.

In section 4.5.4, mode R45 appeared as the most energetic component of the transient response. The effect of increasing its damping is first analyzed using 2, 5 and 10% ratios. A transition towards a new limit cycle of larger brake torque amplitude is observed in figure 5.17.

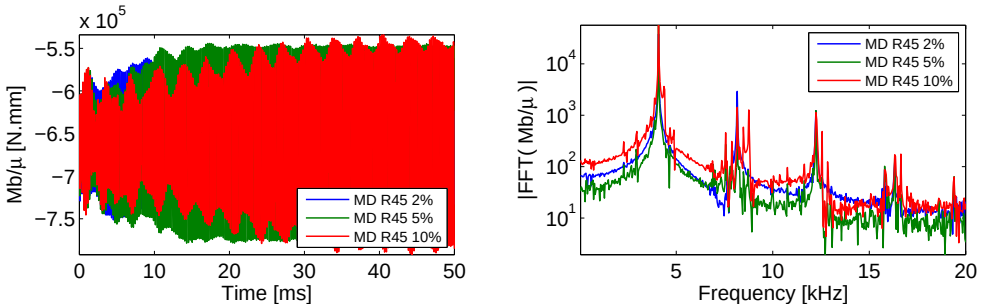


Figure 5.17: Braking torque and its FFT for a limit cycle continuation with mode R45 damping, as function of the ratio

Damping mode R45 thus only does not tame the squeal instability. Although the energy contribution of mode R45 decreases when increasing its damping ratio, mode R44 participation increases and then becomes preponderant in the response for high mode R45 damping ratios. Figure 5.18 plots the modal mechanical energy variations of modes R44 and R45. The contribution of mode R45 is killed by the damping ratio applied, however, a new excitation occurs by mode R44.

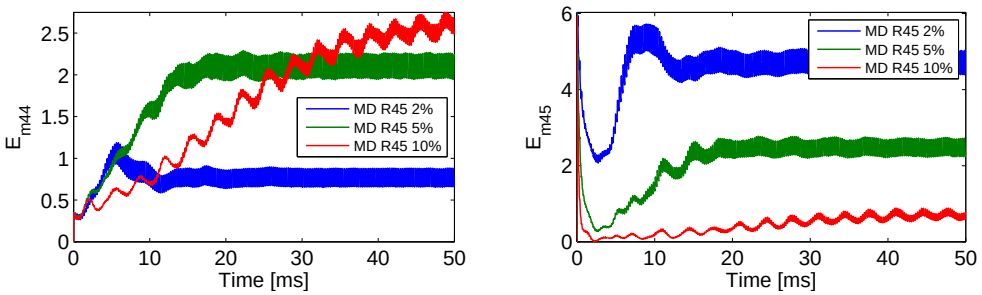


Figure 5.18: Modal sensors R44 and R45 for a limit cycle continuation with mode R45 damping, as function of the ratio

The slip-separation pattern observed in the limit cycle does not disappear when damping mode R45, as plotted in figure 5.19. The actual amplitude of the pseudo separation is decreased but not to significant levels. This can be explained by looking at the modeshapes R44 and R45 (figure 3.22). Damping mode R45 impacts the pad/disc interaction, however, the caliper vibration of mode R44 keeps the pad moving on the disc through another mechanism.

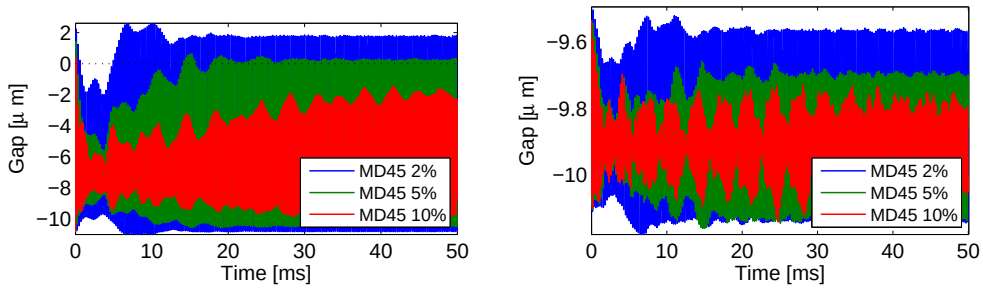


Figure 5.19: Effect of mode R45 damping on a slip-separation pattern, two sample contact points

This shows the importance of mode R44 in the response although it is not energetic in the nominal limit cycle. Once mode R45 is damped, the energy of mode R44 grows as the slip-separation pattern is more energetic. This results in a larger destabilization, observed in the braking torque signal.

The effect on the disc vibration is rather low. Figure 5.20 shows that points vibrating with a large amplitude in the initial limit cycle show similar levels with a slight decrease of the average level. In fact the disc vibration shows more modulation and points with a lower level of vibration show an amplitude increase. Damping mode R45 thus has a small effect. The effect in terms of noise reduction may not be visible.

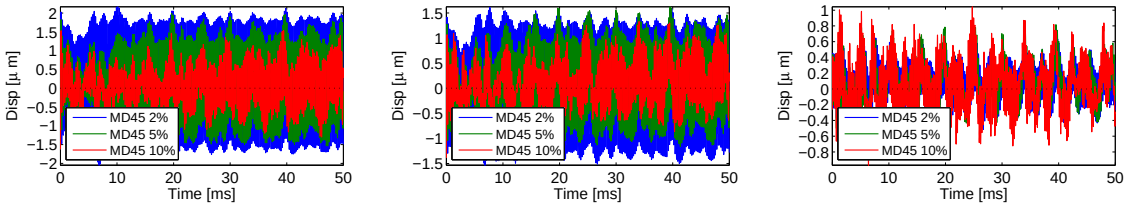


Figure 5.20: Vibration levels around the vertical direction taken on the disc ODS (points 11, 29 and 48) when damping mode R45

Since mode R44 was shown to have a strong influence, its damping is analyzed. Damping mode R44 has a positive effect on the limit cycle since its amplitude decreases, even for relatively low damping ratios, as illustrated in figure 5.21. Large damping ratios however fail to further tame the instability.

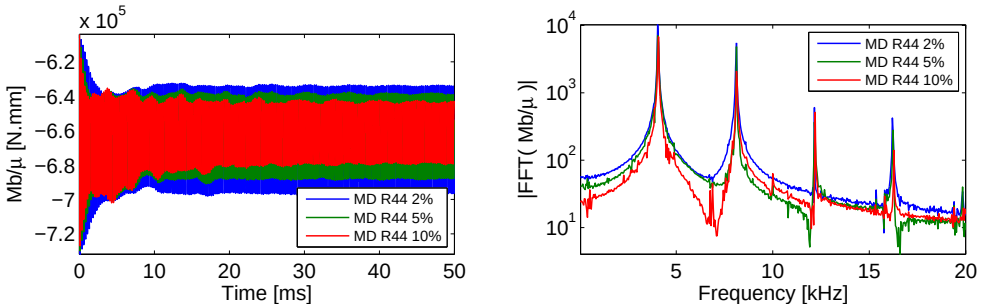


Figure 5.21: Braking torque and its FFT for a limit cycle continuation with mode R44 damping, as function of the ratio

Figure 5.22 shows that a low damping ratio is enough to remove an energetic participation of mode R44 to the response as mode R44 sensor goes to zero. The incident effect on mode R45 is however impressive as its energy decreases significantly as function of mode R44 damping ratio. It is noticeable that the decay rate of mode R45 is constant when R44 damping is increased, thus confirming the coupling between both modes.

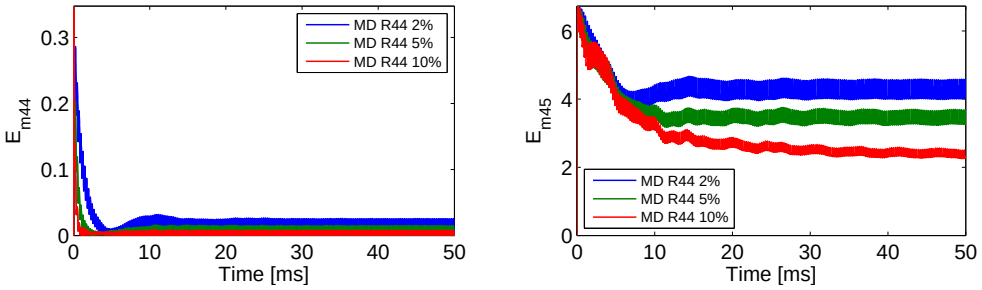


Figure 5.22: Modal sensors R44 and R45 for a limit cycle continuation with mode R44 damping, as function of the ratio

Damping mode R44 has a great impact on the slip-separation pattern. Figure 5.23 shows a sample contact point from the rear of the outer pad (slip-separation area). The amplitude is much decreased than when damping mode R45 directly. Again looking at modeshapes R44 and R45 (figure 3.22) shows that damping the pad/caliper interaction stabilizes the pad excited at the disc level, so that vibrations can effectively be decreased.

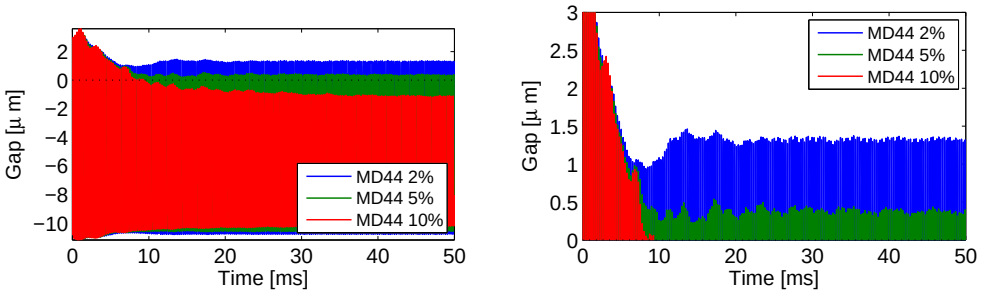


Figure 5.23: Effect of mode R44 damping on a slip-separation pattern, sample contact point. Left: global view. Right: zoom in

The effect of damping mode R44 is good on the disc vibration levels. Figure 5.22 shows however that the unstable mechanism is not altered, although its amplitude is decreased. Vibration levels of the high amplitude points are decreased by 30% from $2.2\mu m$ to $1.5\mu m$ with a 2% damping ratio, as seen in figure 5.24.

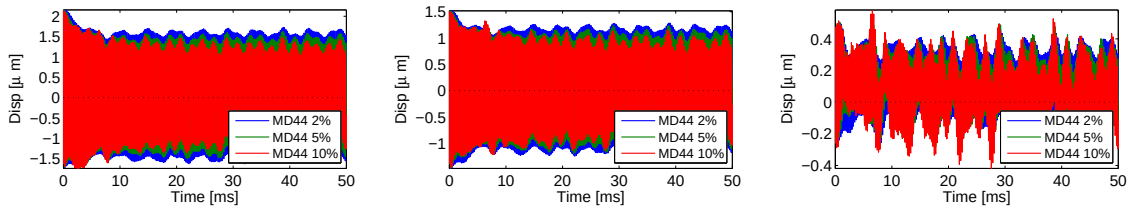


Figure 5.24: Vibration levels around the vertical direction taken on the disc ODS (points 11, 29 and 48) when damping mode R44

As a final case, damping both modes R44 and R45 at a 2% ratio is enough to tame the instability. This is shown in figure 5.25 where the initial limit cycle converges back toward the steady state. These results are similar to the results obtained when recomputing the solution from the steady state.

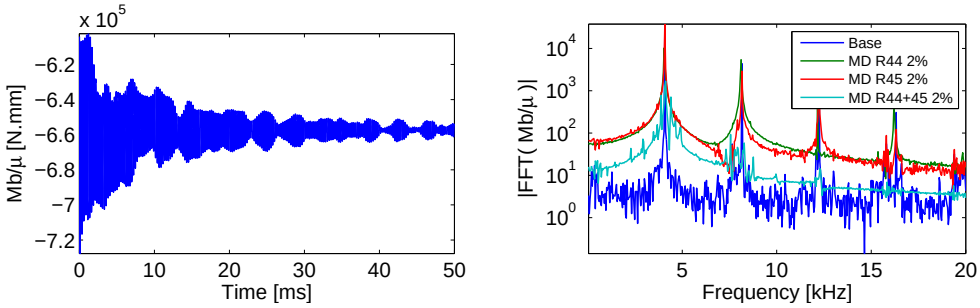


Figure 5.25: Left: Braking torque for a limit cycle continuation with mode R44 and R45 2% damping. Right: braking torque FFT for several modal damping strategies over modes R44 and R45

The interaction between modes R44 and R45 is emphasized in figure 5.26 showing modal sensors R44 and R45 for several damping strategies in a simulation continued from the base limit cycle. Damping mode R45 increases mode R44 participation while mode R45, although damped at the beginning of the transition follows a new divergence phase before transiting back to the new limit cycle. The transient phase is rather disturbed and shows the generation of a potentially different limit cycle. Figure 5.27 eventually shows that damping mode R45 has a low effect on the slip-separation pattern observed.

Damping mode R44 kills its response in the limit cycle and decreases mode R45 to comparable levels than when directly damped. The transition to a new limit cycle is rather quick and simple to assess as the level of energy diminishes following an exponential decay. The effect on the stick-separation patterns plotted in figure 5.27 is greater than when damping mode R45, which should explain the global energy decrease observed. This is compatible with modeshapes R44 and R45.

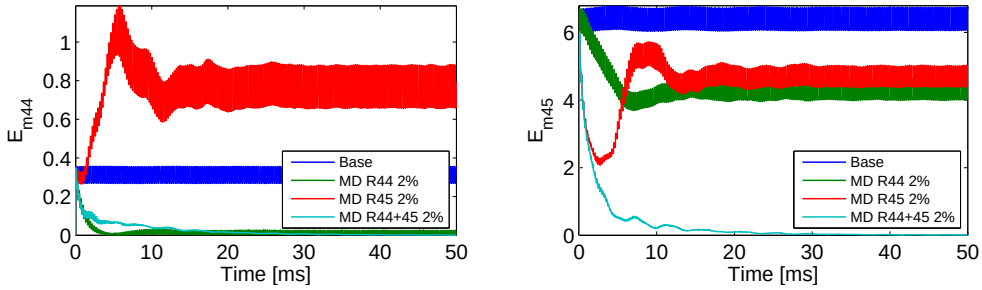


Figure 5.26: Modal sensors R44 and R45 for several modal damping strategies over modes R44 and R45

When both modes are damped, both modes follow the same initial decay energy than when only one of them is damped. No interaction is however possible in this case, such that the limit cycle disappears. The slip-separation pattern in fact instantly disappears in this configuration, figure 5.27.

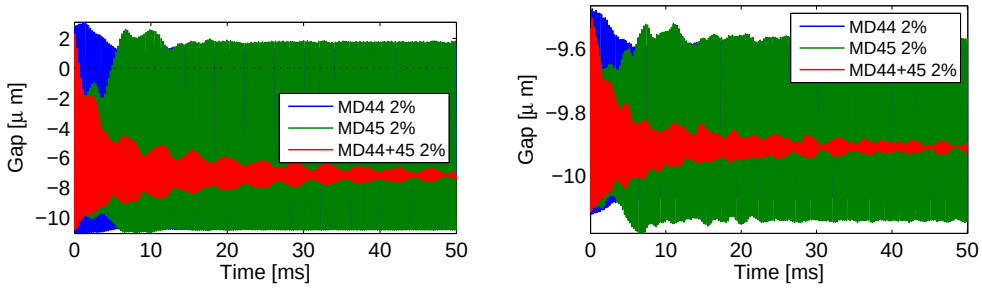


Figure 5.27: Effect of mode R44+R45 damping on a slip-separation pattern, sample contact points

The disc vibration level shows the same tendency. It can be seen in figure 5.28 that damping mode R44 is more efficient than damping mode R45, while damping both is enough to stop the vibration.

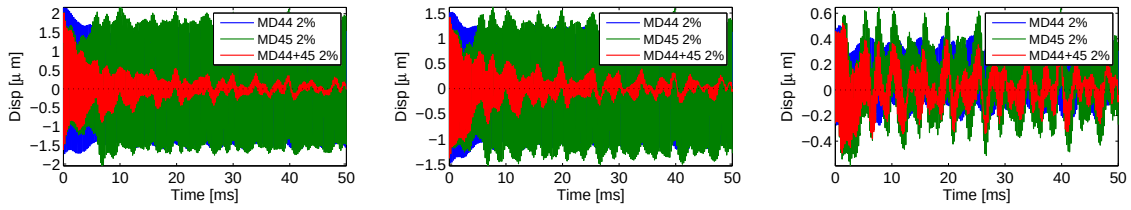


Figure 5.28: Vibration levels around the vertical direction taken on the disc ODS (points 11, 29 and 48) when damping modes R44 and R45

Using modal damping on either mode R44 or mode R45 makes the system converge towards a different limit cycle than the initial one. Figure 5.29 illustrates this by showing the limit cycle obtained in the different cases. The amplitude increase when damping mode R45 and decrease when damping mode R44 are clearly observed patterns.

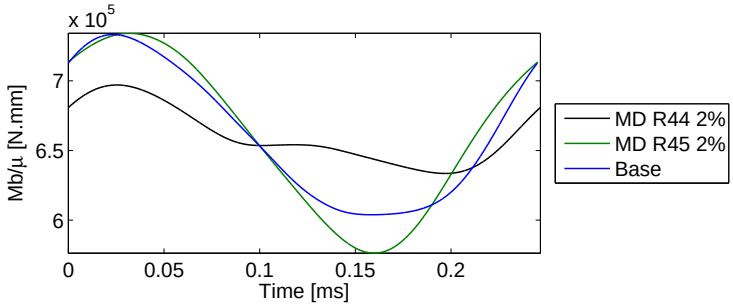


Figure 5.29: Identification of the new limit cycles from the continuation of the base computation with a 2% damping of either mode R44 or mode R45

Although looking different regarding the braking torque pattern, the cycle does not really change in characteristics. A space-time decomposition based on the SVD is performed for each cycle plotted in figure 5.30. MAC computation are performed between each set of principal deformation in figure 5.31 to quantify the differences.

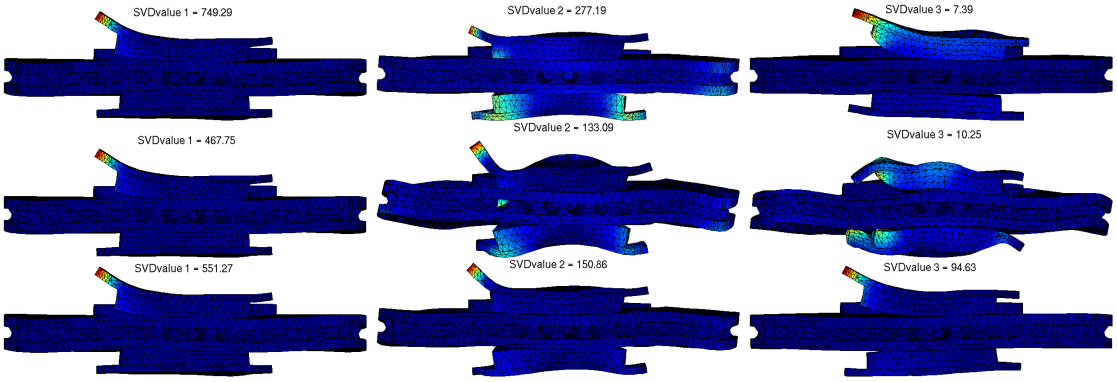


Figure 5.30: Principal shapes extracted from the limit cycle continuation. Top: nominal (no modal damping) simulation. Middle: 2% mode R44 damping. Bottom: 2% mode R45 damping

The first singular shape is identical for all cycles, showing a MAC of almost 100% between all shapes. The singular value obtained is different for each cycle however, as it depends on the global amplitude of the cycle. Since the first singular shape is correlated to mode R45 (74% MAC), its amplitude decreases for both damping strategies, as observed in the previous section.

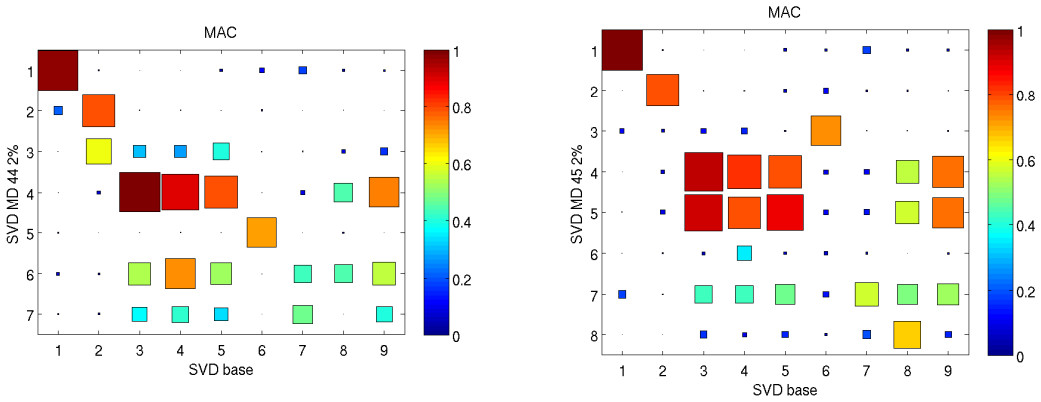


Figure 5.31: MAC between the SVD shapes extracted from the limit cycle continuation with 2% mode R44 damping and 2% mode R45 damping in comparison to the base cycle

The second shape does not differ between the base computation and the case with mode R44 damping. This is coherent with previous observations. The total amplitude of the response is lower in this case and is observed by the singular values.

More differences occur between the base computation and the case with mode R45 damping. In particular, a third significant dimension is observed in figure 5.30, which is absent or insignificant in other cycles, looking at the singular values of figure 5.30. The limit cycle in this case is richer, which was expected due to the transitions observed in the previous sections.

Figure 5.32 eventually presents the cross MAC between both new cycles. Although both second singular vectors correlate reasonably with the second singular vector of the initial cycle, their correlation is quite low thus indicating a significant shape change.

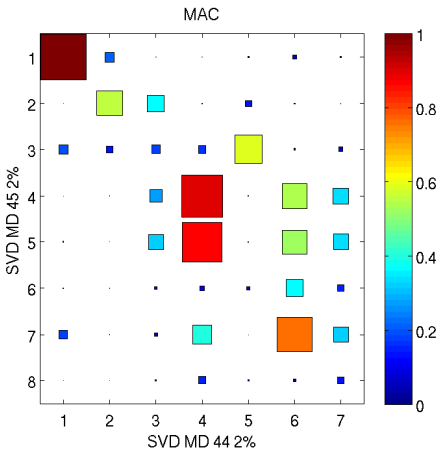


Figure 5.32: MAC between the SVD shapes extracted from the limit cycle continuation with 2% mode R44 damping and 2% mode R45 damping

In this section, it was observed that adding damping does not always have a positive effect. Indeed, applying modal damping to mode R45 actually destabilizes the system, as the limit cycle amplitude increases. This is a well known fact and similar results were shown by Fritz *et .al* [168] and Hervé [169]. Similar patterns were also obtained experimentally by Cantone [100].

The criterion retained in these references to decide whether damping can be guaranteed to have a positive effect relies on the concept of *proportional damping*, as presented by Sinou *et al.* [170]. The study then proves that if damping in increased in identical proportions on the whole system, it will have a positive effect, while non-proportional cases can show peculiar effects.

Physically, damping is enhanced by devices dissipating energy by viscoelastic behavior or friction. Such devices are localized in space so that the mathematical assumption of proportional damping is not verified and detailed studies such as those shown here will be needed.

The observations concerning the modal interactions can be understood looking at the shapes of modes R44 and R45, presented in section 3.4.2. The phenomenon observed can be repeated in the frequency domain, by applying the corresponding modal damping matrix – trivial in the frequency domain – to the complex mode computation. Figure 5.33 plots the predictions obtained, which are in good correlation with the observations.

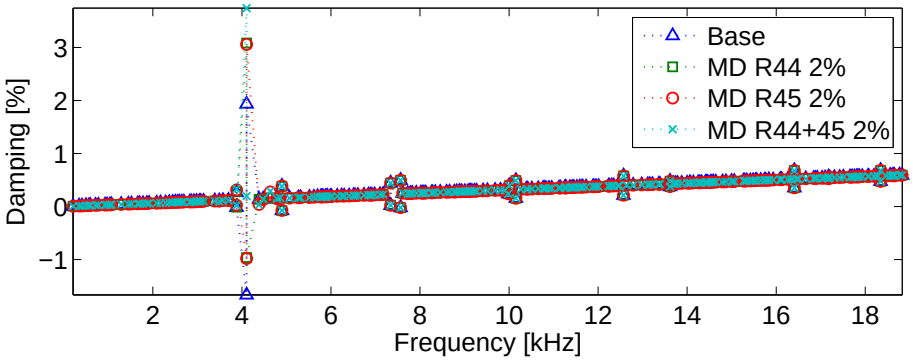


Figure 5.33: Industrial brake stability diagram at 12 Bar as function of real modes R44 and R45 modal damping

The information is altered as the same real part is obtained by damping either mode R44 or R45, although the transient result is different. The stabilization by damping both modes is however predicted.

Considering both frequency domain computations for quick analyzes and time domain predictions for confirmation and estimation of limit cycle amplitudes thus seems the appropriate strategy to retain in future design work.

5.2 Using component properties in design studies

The damping results of section 5.1.2 provided a better understanding of how modal interactions occur in a limit cycle. The exploitation in terms of design is however still difficult, as stated in chapter 2. Working at the component level, similar studies as previously presented can be performed. The application of the CMT method is thus tested for stability predictions, section 5.2.1.

As observed in section 5.1.2, the frequency domain predicted a decrease of mode C44 damping ratio when damping mode R44 or mode R45. The impact on the initial growth rate of the mode – which corresponds to the modal damping ratio – does however not provide any information on the limit cycle amplitude. In particular it was shown that damping mode R45 only increased the limit cycle amplitude – although the growth rate was decreased. Time integration is therefore needed for validation purposes, and is the object of section 5.2.3.

The application of damping is more practical in the time domain than applying a stiffness variation. The initial equilibrium is indeed not altered. Besides, damping treatment of industrial components is a well-known approach at SDTools and has been the object of the PhD of Groult [26]. Such orientation is therefore sought in the present work, where component modal damping will be tested, and serve as a validation mean in the time domain.

5.2.1 Sensitivity and frequency reanalyzes for component stiffness changes

The three most unstable modes of the industrial brake at 12 Bar are plotted in figure 5.34. The component-wise strain energy ranking gives information concerning the component participation. Much more information can however be obtained by the CMT method.

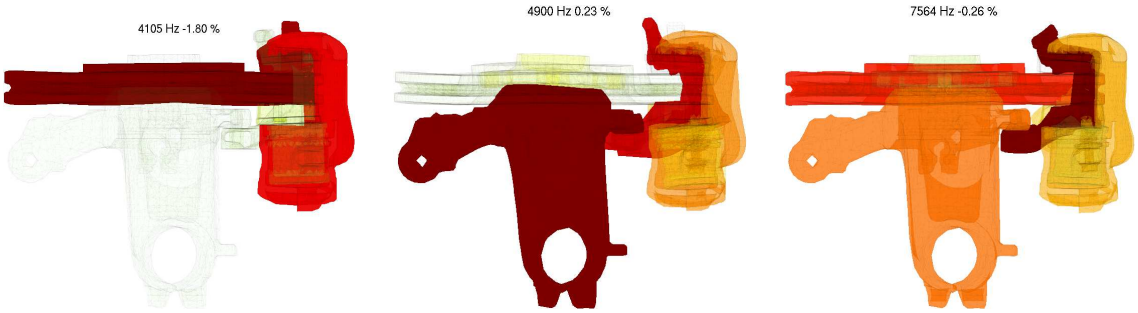


Figure 5.34: Main unstable modes at 12 Bar, C44, C51 and C84, strain energy plotted and ranked by components

In a CMT model, the sensitivity of the real part of a complex mode to a specific component mode is easily computed using equation (2.26). The complete evaluation for all assembled modes as function of all component modes is thus obtained in only 60s CPU. Results are shown in figure 5.35.

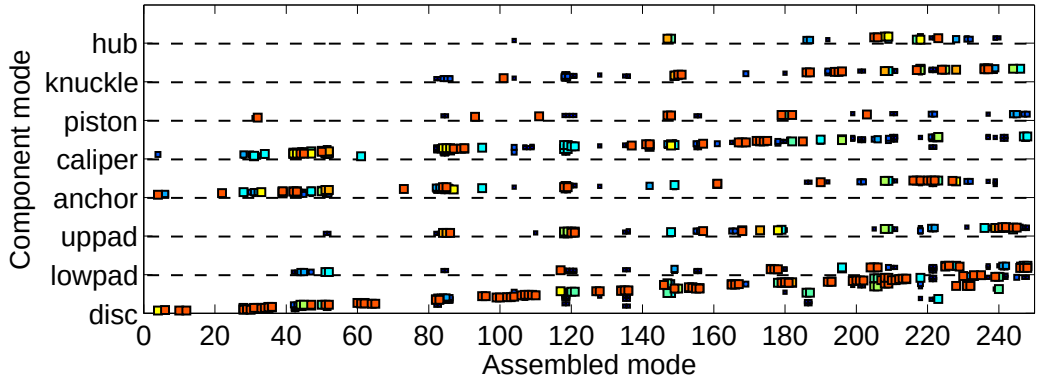


Figure 5.35: Real part sensitivity of the first 250 assembled modes as function of all retained component modes. The hotter the color, the larger the marker, the higher the sensitivity

The first visible effect is that a number of stable modes can become unstable for certain component mode frequency variations. It can be observed that the biggest role players are the disc, caliper and anchor. Surprisingly, the pads do not show a great influence on the stability. Regarding the nominally unstable modes, although their strain energy distribution by component is different, figure 5.36 shows the components having an influence are the same.

The plots of figure 5.36 are obtained by plotting vertical lines of the global picture of figure 5.35. The caliper appears as a great instability player for all three modes. The disc is also always present. Pads are eventually present but to a lower scale.

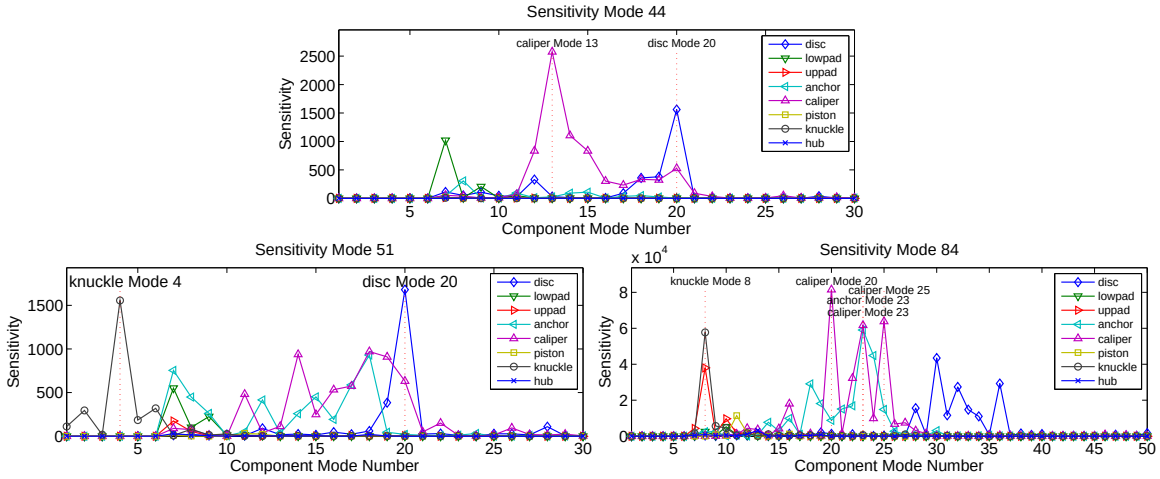


Figure 5.36: Unstable modes sensitivity to component modes (eigenvalue real part sensitivity). Zoom ins for the 3 most unstable modes

The scaling has been chosen arbitrarily to obtain significant sensitivities, the only relevant information is the relative ranking. The validity range of the value is however small and not controlled. Actual prediction will be performed in the following using re-analysis, which is much more accurate, as observed by Bobillot [41].

Mode C44 has been observed as the main cause of instability in both time and frequency domains. The remaining of this section will thus focus on analyzing a modification altering the stability of this mode. From figure 5.36, mode C44 is sensitive to modes of three components: the disc (mode 20), outer pad (mode 7) and the caliper (modes 12 to 14).

Analyzing each influent component mode separately allows some observation in terms of strain energy localization, as plotted in figure 5.37. The disc mode is radial with 5 lobes showing a lobe distance of roughly the pad length. Most strain energy is located in the ventilation columns. The pad mode is the first bending mode, driven by the backplate which strains the lining. The caliper modes show high strain energy at the fingers root, which is the result of large displacements at the pad/caliper interaction area.

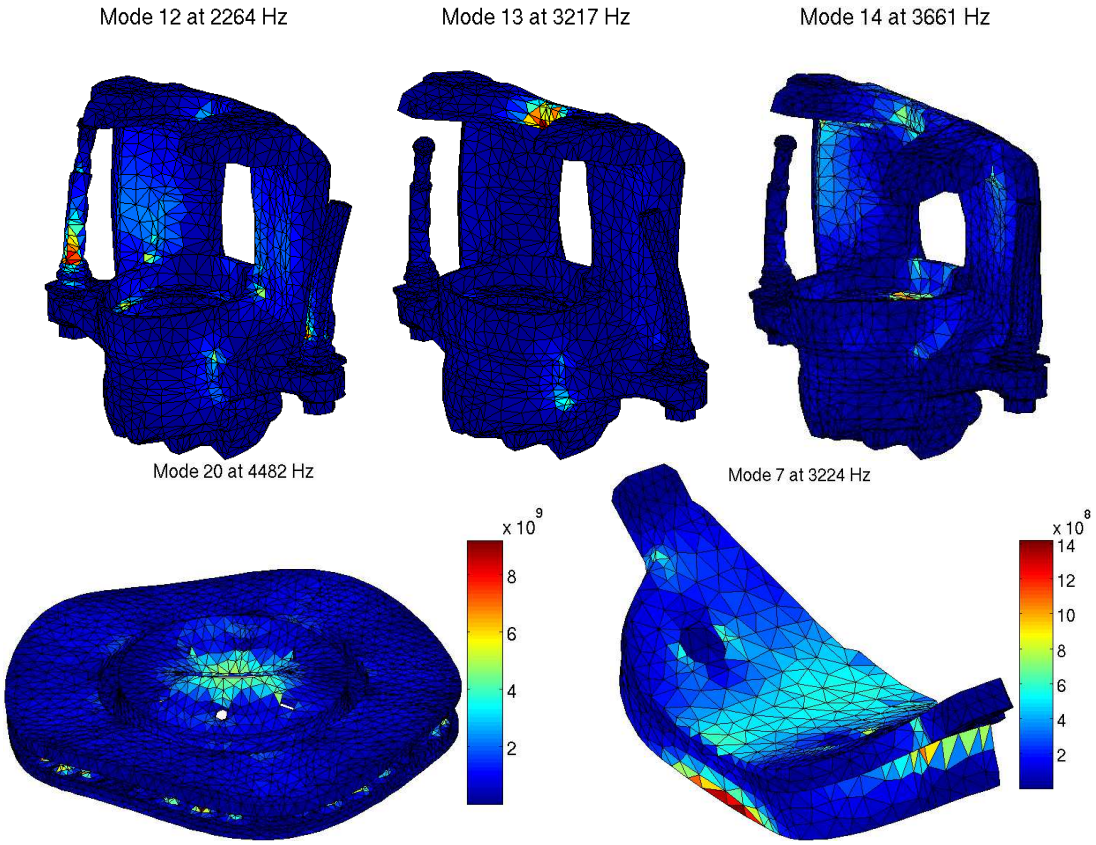


Figure 5.37: Strain energy localization of the three most influent component modes, on unstable mode C44

The relevance of this information is tested against observations in the time domain. The nominal singular shapes obtained in section 4.5.4 can be characterized on the component mode model for correlation purposes.

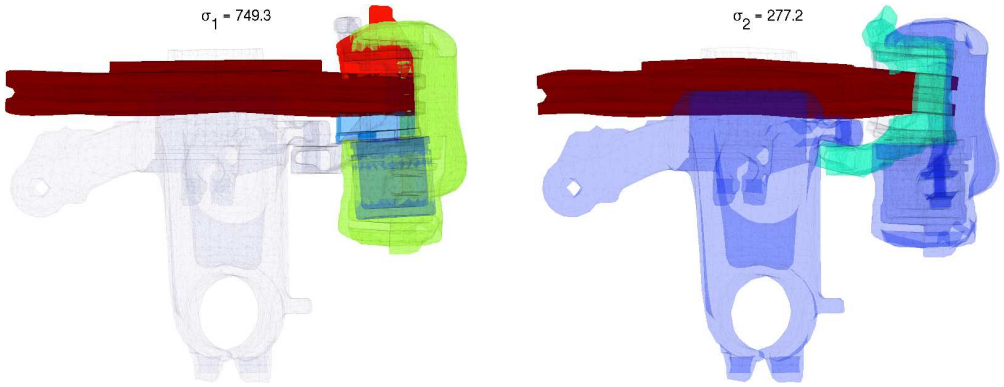


Figure 5.38: Projection of the main shapes computed by the SVD on the component mode model. Strain energy distribution by component

Figure 5.38 shows the distribution of energy by component and should be viewed in parallel of figure 5.34. The energy rankings are quite different as the principal shapes show less energy in the caliper relatively to the disc and outer pad than for mode C44. This is coherent regarding the difference existing between the projection of mode C44 on the real modes, figure 3.22 and the modal sensor results of figure 4.53. Mode R44 is then less present in the time response than *a priori* computed for mode C44.

The singular vector expression in the component mode model coordinates directly gives the relative participation of each component mode to the response. The values are plotted in figure 5.39 for the first two singular shapes in amplitude.

Similar observations than from the sensitivity analysis can be made. The highest value in figure 5.39 corresponds to a rigid body mode, which are present in the CMT. Pad mode #7, disc mode #20 and caliper modes #12 to #14 are nevertheless clearly highlighted.

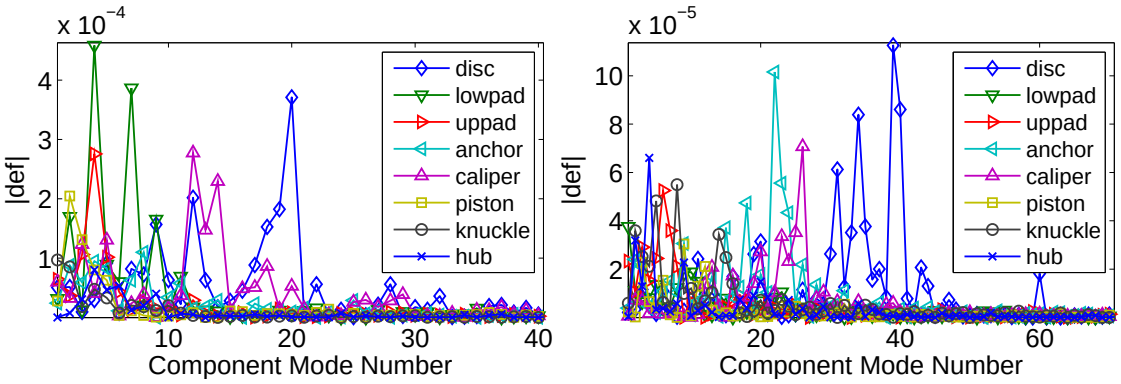


Figure 5.39: Projection of the principal shapes given by the SVD on the component modes through the component mode model. Left: SVD singular shape #1. Right: SVD singular shape #2

The second singular shape mainly corresponds to mode C95, close to the second harmonic of mode C44 at 8kHz. It can be seen that the disc modes participating for the first singular shape at 4kHz is mode 20, while the disc modes participating for the second singular shape at 8kHz is modes 40. The interest of this data is limited, since the second singular shape corresponds to the forced response of mode R95.

Trying to work on the disc or pad is not easy in real life applications, since these components are constrained by performance requirements, and are not always fully produced by Bosch. The results presented here show that in addition, it is not the most relevant thing to do. Figure 5.40 shows that trying to change disc mode frequencies does not improve significantly mode C44 stability. Similar effect is obtained when trying to alter the outer pad first bending modes (mode 7, 9 and 11), as presented in figure 5.40.

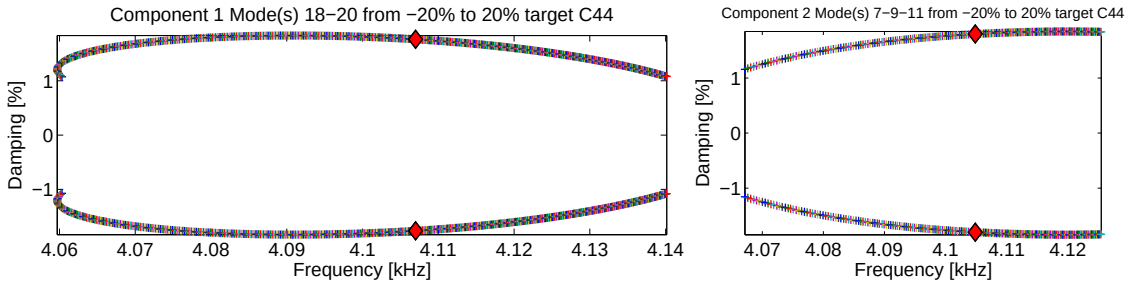


Figure 5.40: Attempt to alter mode C44 instability by tuning disc modes 18 and 20 (left) and by tuning the first pad bending modes (7, 9, 11) (right). Red diamond stands for the nominal state

Figure 5.41 shows that altering caliper modes is more effective. Tuning caliper mode 13 only is maybe not physically possible, but shows that mode C44 can be fully stabilized. The counter-part is the apparition of an unstable mode C42 for rather low frequency variations. In fact the stability does not really disappear but is transferred to mode C42.

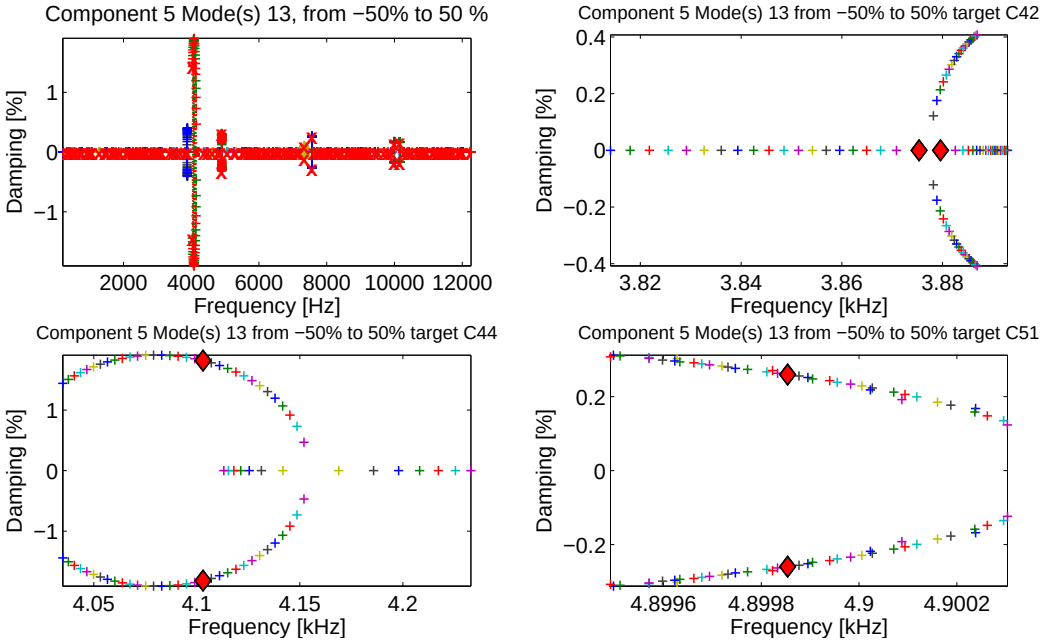


Figure 5.41: Effect of caliper mode 13 frequency tuning from -50 to +50 % on the global system stability at 12 Bar. Red diamond stands for the nominal state

Altering caliper modes 12 to 14 altogether seems more physical as their displacement patterns are close (movements of the caliper fingers). Although caliper mode 13 is the most sensitive, modes 12 and 14 are not fully negligible in figure 5.36. The results of such modification are presented in figure 5.42. Mode C44 does not become stable, but no transfer to mode C42 is observed. The effect on mode C51 is also limited.

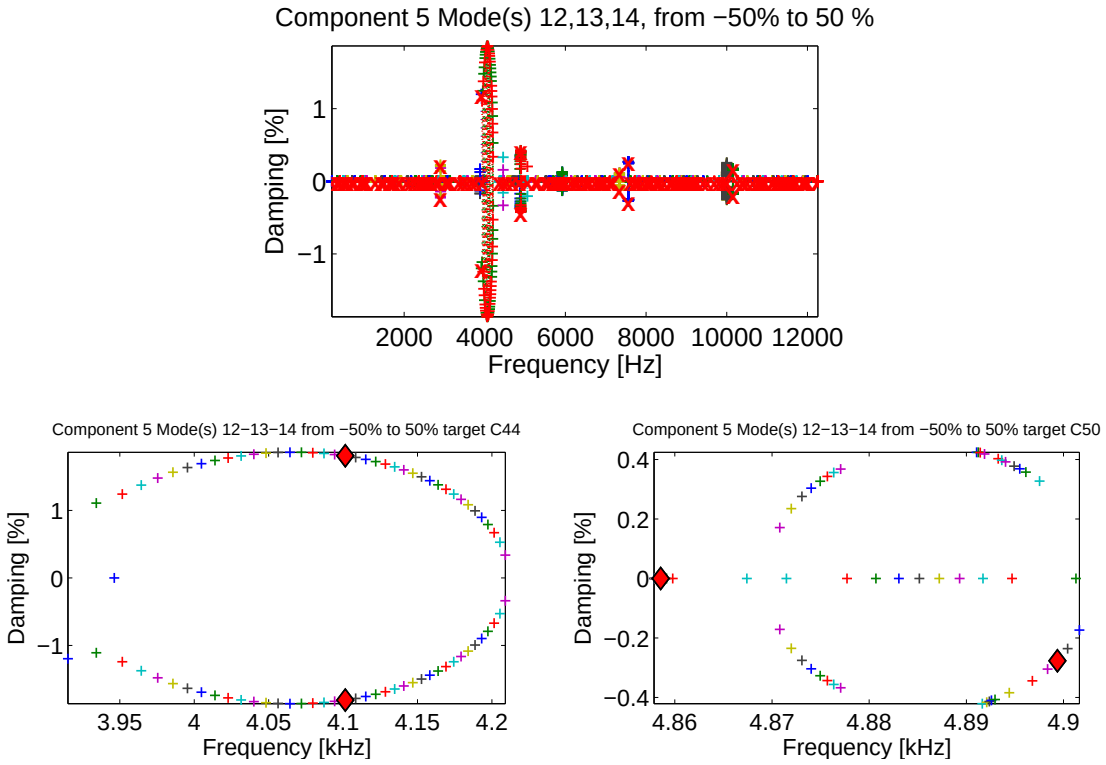


Figure 5.42: Effect of caliper modes 12, 13 and 14 frequency tuning from -50 to +50 % on the global system stability at 12 Bar. Red diamond stands for the nominal state

The CMT study on stiffness thus points the caliper as being the most relevant component to modify. Global modal damping studies of section 5.1.2 pointed mode R44, and the decomposition of mode C44 on real modes in figure 3.22 showed that the caliper is most energetic for mode R44. The results found here are thus consistent with earlier analyzes.

5.2.2 Component damping modification

Working in the component reduction basis (CMT), each component mode damping ratio appears on the diagonal of the damping matrix, whose expression is

$$[C_{Comp}] = \left[\backslash 2\zeta_{ckl}\omega_{ckl}\backslash\right] \tag{5.3}$$

where $[C_{Comp}]$ is the damping matrix projected in the component mode basis, ζ_{ckl} is the damping ratio associated to component k mode l. Parametric analyzes similar to those of the previous section can be performed and are illustrated in the following.

Trying to damp disc modes or pad modes does not yield promising results. In particular, the damping evolution of mode C44 shows inflections that do not converge towards stability. This can be observed in figure 5.43 for the disc tuning and figure 5.44 for the pad tuning.

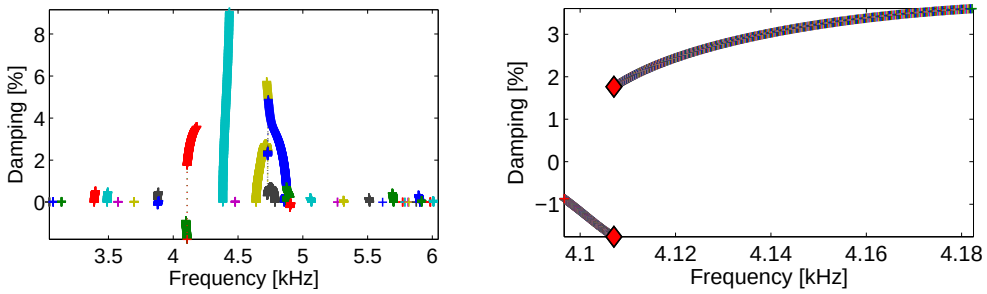


Figure 5.43: Effect of damping disc modes 18 and 20 from 0 (red diamond) to 20% on the stability of the global system, and zoom in on mode C44

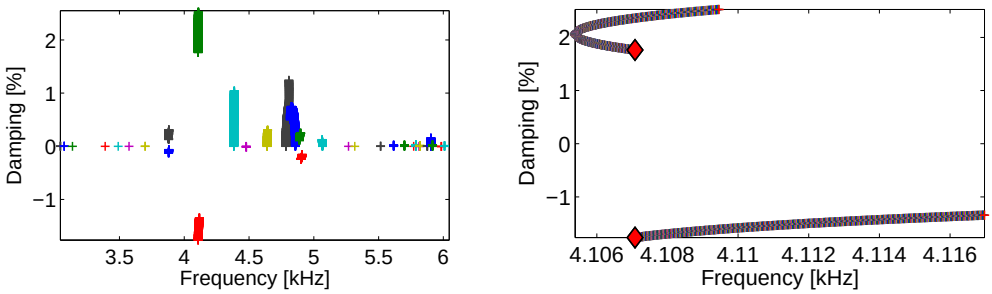


Figure 5.44: Effect of damping outer pad mode 7 from 0 to 20% on the stability of the global system and zoom in on mode C44

Damping caliper modes show again better results, with the possibility to stabilize mode C44. Figure 5.45 shows global damping evolutions obtained when altering caliper mode 13 damping. A lot of unstable modes are improved from the modification.

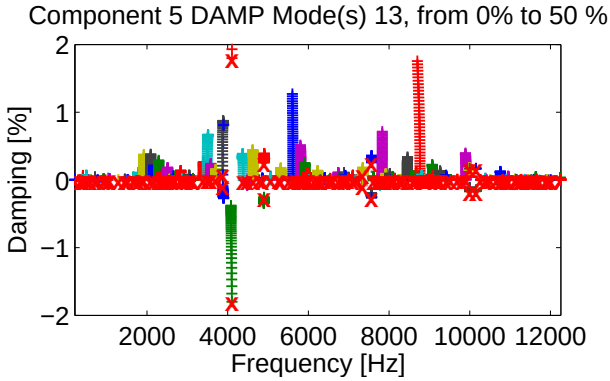


Figure 5.45: Effect of damping caliper mode 13 from 0 to 50% on the stability of the global system

Figure 5.46 shows details from figure 5.45, in which the instability transfer from mode C44 to mode C42 is observed. The instability does not disappear completely in this case. Besides, mode C51 is also destabilized.

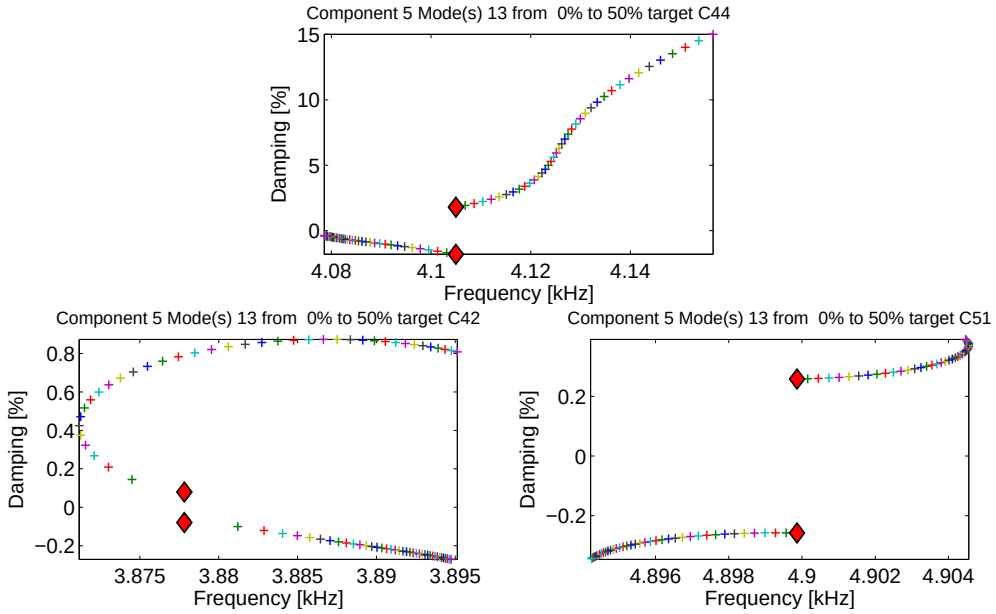


Figure 5.46: Effect of damping caliper mode 13 from 0 (red diamond) to 50% on the stability of the global system, zoomed in on system modes C44, C42 and C51

Damping caliper modes 12 to 14 seems a better option. Figure 5.47 shows that mode C44 is stabilized and that the instability is not transferred to another close mode. Mode C51 is however quickly destabilized, which may be an issue.

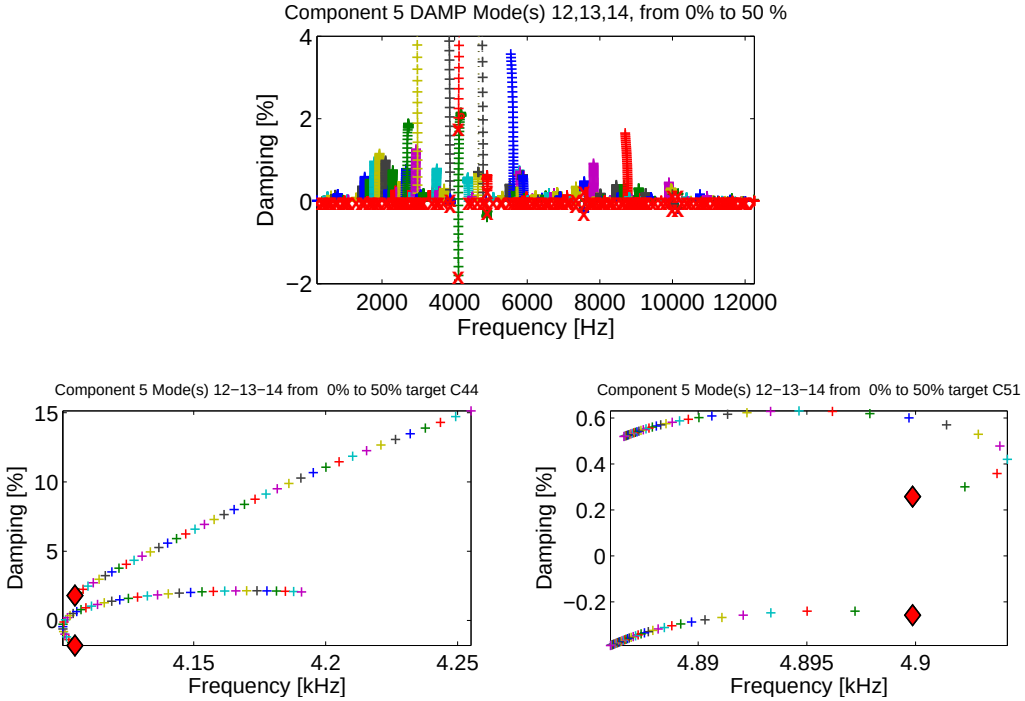


Figure 5.47: Effect of damping caliper mode 12 to 14 on the stability of the global system, from 0 (red diamond) to 50%. Zooms in are performed for modes C44 and C51

The damping treatment study shows the same trends than the component mode frequency tuning. The component interaction responsible for the instability is mostly due to the caliper. The study also highlighted the fact that such component tuning can destabilize other modes. A reevaluation of results is thus needed and time domain analyzes will provide the mean to validate if other instabilities start to appear.

5.2.3 Time domain damping of component modes

The objective of this section will be to test component mode damping in the time domain. The simulations were however performed using a model reduced on the system modes (see section 2.2.5). As a first step, the ability to still approximate component modal damping is evaluated.

For component k mode l, namely ϕ_{ckl}, its modal coordinates $\tilde{\phi}_{ckl}$ on the reduced part of the time model is obtained by computing the modal projection on the reduced part of the assembled brake $(as|ref)$,

$$\tilde{\phi}_{ckl} = \left(\left[M_{as|red} \right] \left[\Phi_{as|red} \right] \right)^T \phi_{ckl} \tag{5.4}$$

where $\left[\Phi_{as|red} \right]$ is the trace of the assembled real modes on the reduced part from the reduction technique presented in section 2.2.5 and $\left[M_{as|red} \right]$ the restriction of the assembled mass matrix to the reduced part. Since the caliper does not have retained DOF in the reduced model, this is sufficient to obtain $\left[\Phi_{as|red} \right]$, the trace of the assembled real modes on the reduced part from the reduction technique.

Since the free displacement of the caliper is constrained by the system coupling, the equivalent frequencies obtained by studying the projected component mode shape on the reduced time model is rather high, as seen in figure 5.48. This is due to the fact that contact interactions are taken into account.

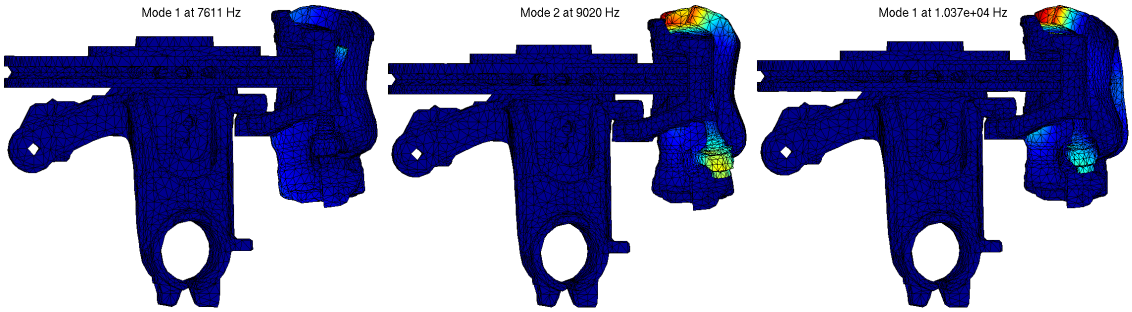


Figure 5.48: Caliper mode shapes 12, 13 and 14 projected on the global time model

Figure 5.48 shows the caliper mode shapes 12 to 14 projected on the reduced time model. The shapes are mainly concentrated on the caliper and seem relevant. In details, figure 5.49 shows that due to the coupling information from the reduction basis, the caliper movement generates some limited pad movement.

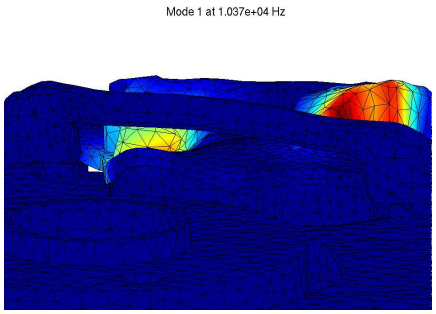


Figure 5.49: Effect of the projection approximation of caliper modes on the time model for caliper mode 13. Color code in displacement

The information loss during the projection is measured by the MAC, in figure 5.50 left. The original caliper modeshapes are compared to the modeshapes projected on the time reduced model and restituted from the assembled real modes reduction basis. The MAC show very good results, which validates the component modal damping strategy in the time domain.

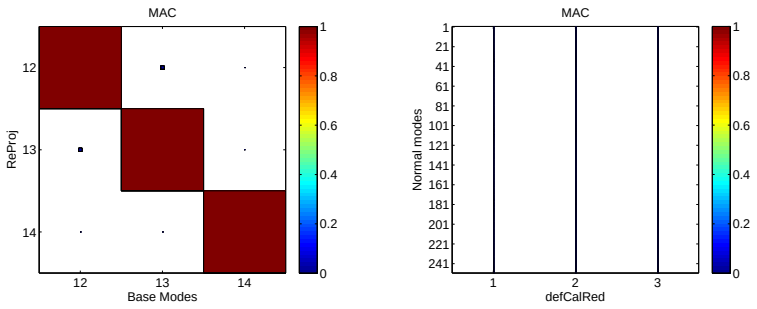


Figure 5.50: Left: MAC between the caliper component modes and the restitution of the projected shapes on the reduced time model restricted to the caliper. Right: MAC between the projected caliper modes and the system real modes

To ensure that the modes projected differ from the real modes, a MAC is also computed between the projected caliper mode shapes and the reduced real mode shapes, figure 5.50right. No correlation exists between the two, which proves the projection does not raise aliasing issues, where a component modeshape could have been identified to a real mode.

The model behavior when damping only caliper mode 13 is first tested. Figure 5.51 plots the braking torque result of this simulation. It can be clearly seen that the limit cycle amplitude is decreased, although no stabilization pattern is observed, as forecast by the CMT analysis. The vibration levels plotted in figure 5.52 are reduced to similar levels than when damping mode R44 in section 5.1.2.

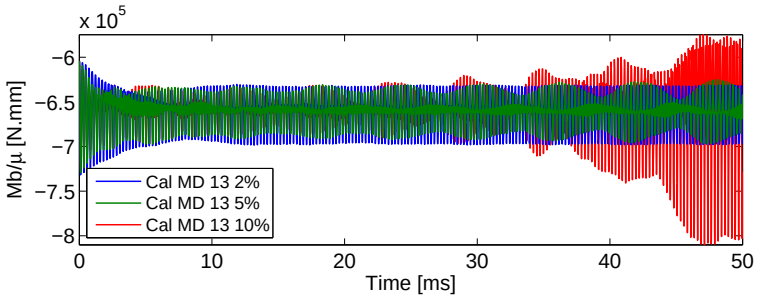


Figure 5.51: Effect of damping caliper mode 13 on the limit cycle obtained at 12 Bar, braking torque

Sample ODS sensors in figure 5.52 show some reduction of the vibration amplitude for the points showing large displacements. A new disc displacement regime can however be observed as the low amplitude point shows a great increase when the caliper mode 13 is overdamped.

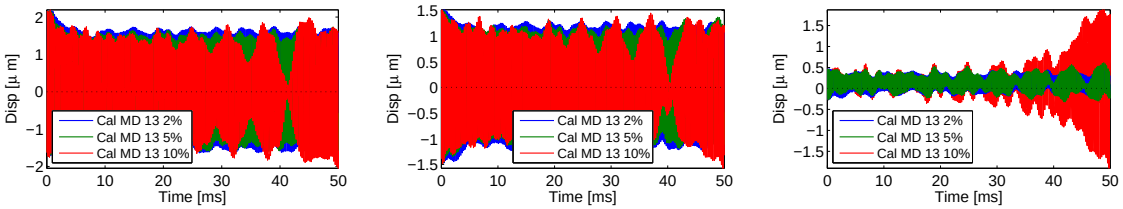


Figure 5.52: Effect of damping caliper mode 13 on the limit cycle obtained at 12 Bar – disc ODS measurements (points 11, 29 and 48)

A new instability clearly appears for large damping ratios, which can be analyzed using modal sensors in figure 5.53. Although modes R44 and R45 contributions are well decreased, mode R42 contribution shows a steep increase, which appears to excite other modes, R43, R44 and R45 afterwards. These results are coherent with the frequency domain study.

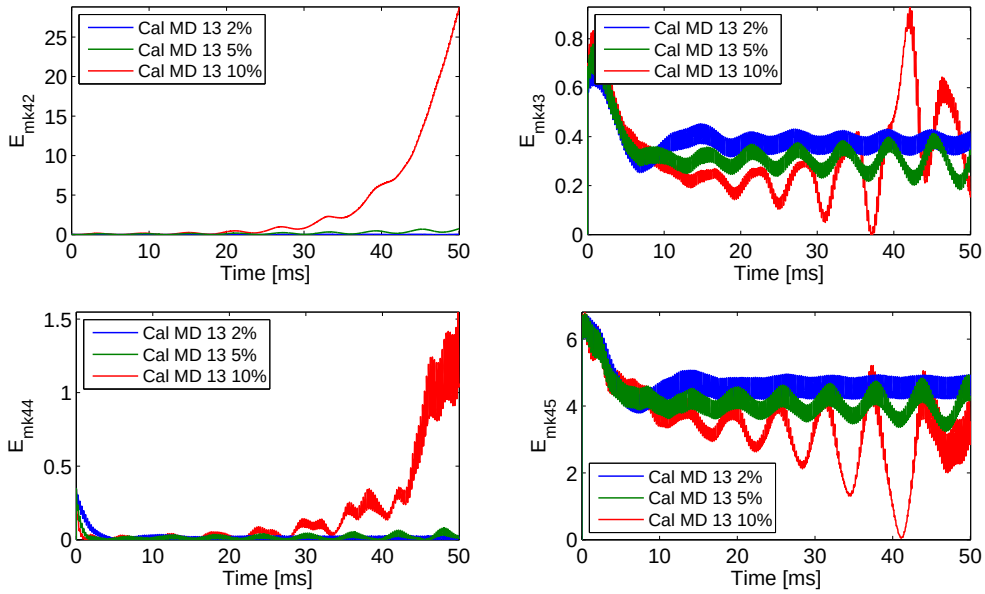


Figure 5.53: Effect of damping caliper mode 13 on the limit cycle obtained at 12 Bar, modal mechanical energies associated to mode C44

The mechanical energy contribution of caliper mode 13 itself is plotted in figure 5.54. A great decay is first observed, a new destabilization in this direction however occurs if the damping ratio becomes too high.

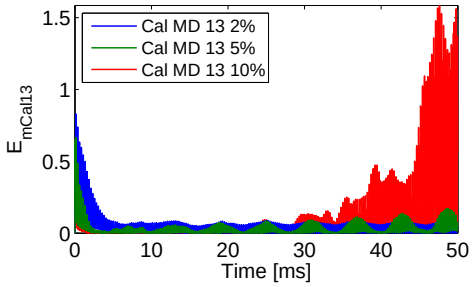


Figure 5.54: Effect of damping caliper mode 13 on the limit cycle obtained at 12 Bar, modal mechanical energy associated to caliper mode 13

The effect on the slip-separation pattern is not fully satisfying as the gap results of figure 5.55 show no real improvement. Damping caliper mode 13 is not a solution in itself. The cycle amplitude decrease is only limited and instability transfers occur, yielding cycles of greater amplitude.

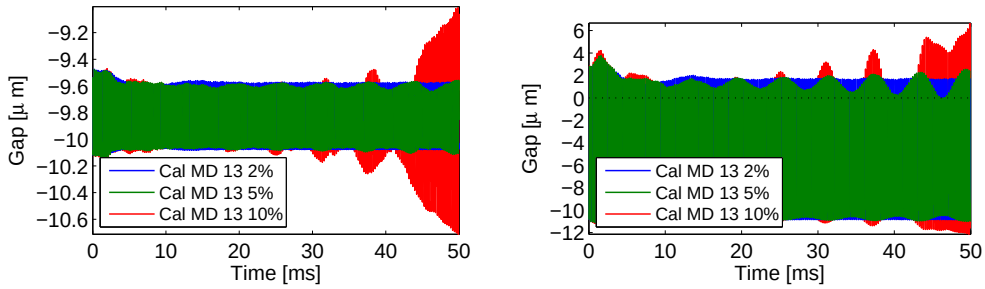


Figure 5.55: Effect of damping caliper mode 13 on the limit cycle obtained at 12 Bar, slip-separation pattern evolution through sample gap measurements

Damping caliper modes 12, 13 and 14 showed better results from the CMT and is tested below. Figure 5.56 shows greater amplitude decrease; although other instabilities seem to occur in the end of the simulation. Their growth rate is also small.

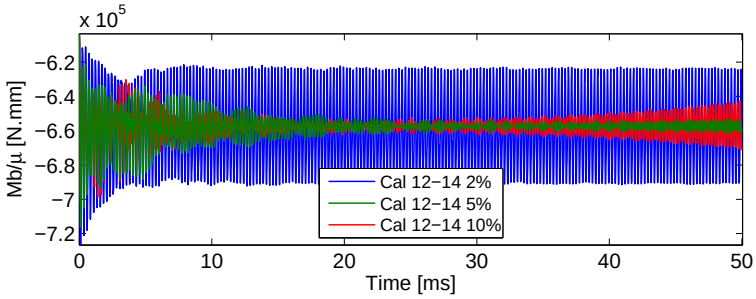


Figure 5.56: Effect of damping caliper modes 12 to 14 on the limit cycle obtained at 12 Bar, braking torque

The displacement results from figure 5.57 show that the vibration levels are dramatically decreased with such treatment. A mild increase occurs in the end which would require follow up simulations to assess the vibration levels attained by the new vibration regime. This is a case of modal competition that will be studied later on.

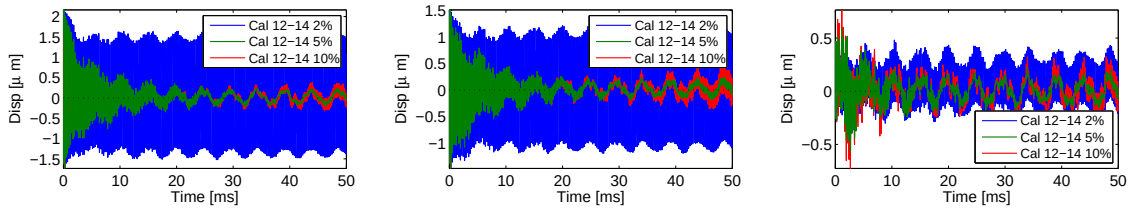


Figure 5.57: Effect of damping caliper modes 12 to 14 on the limit cycle obtained at 12 Bar, disc ODS measurements (points 11, 29 and 48)

In figure 5.58, the modal mechanical energies of modes R42 to R45 which contribute to the cycle of mode C44 only show global energy decreases. Some initial disturbance is observed for mode R42, as when damping only caliper mode 13. No global instability transfer however results from the disturbance, since caliper modes 12 to 14 are now also damped.

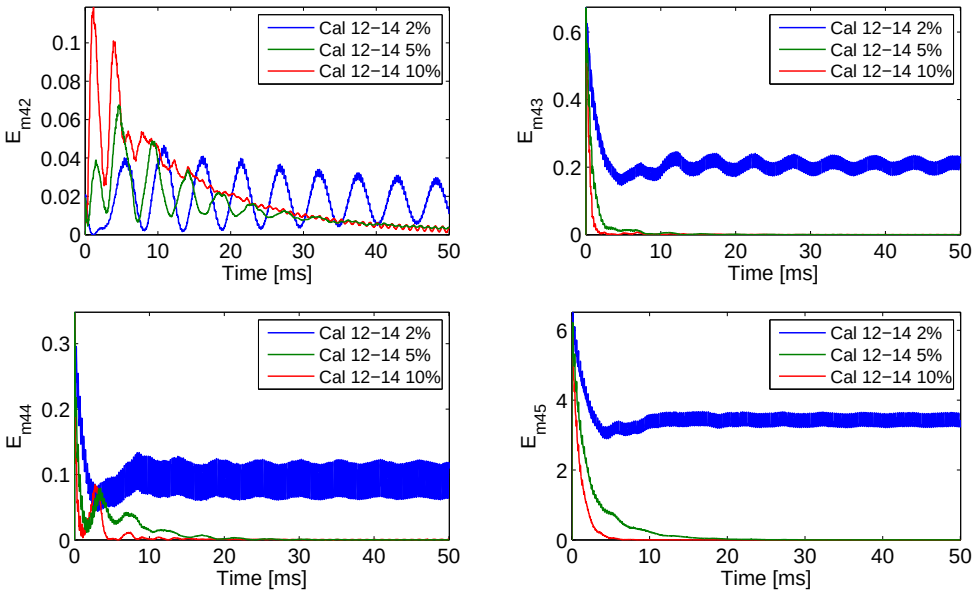


Figure 5.58: Effect of damping caliper mode 12 to 14 on the limit cycle obtained at 12 Bar, modal mechanical energies associated to mode C44

Applying a 2% damping ratio to caliper modes 12 to 14 does not change the instability regime, which can be seen on all curves of this section. Figure 5.59 shows in particular that the contribution from the caliper modes to the response are not decreased to zero in such configuration. Applying a higher damping ratio however kills the caliper contribution to the response, which validates the observed change in instability regime.

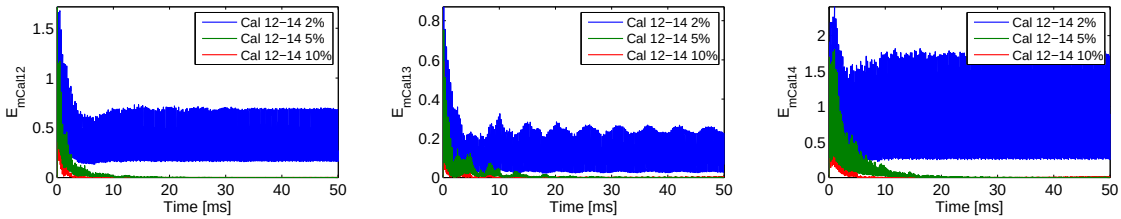


Figure 5.59: Effect of damping caliper modes 12 to 14 on the limit cycle obtained at 12 Bar, modal mechanical energy associated to caliper modes 12, 13 and 14

As verified in section 4.5.2 analyzing modal competition, another unstable mode starts to grow once the current instability is killed. Figure 5.60 indeed shows that mode C51, through its main contributor mode R51 starts to show an exponential growth. The logarithmic plot allows seeing that the growth in fact occurs very early. Consistently with the CMT analysis, the higher the damping ratio applied to the caliper modes, the higher the instability ratio of mode C51.

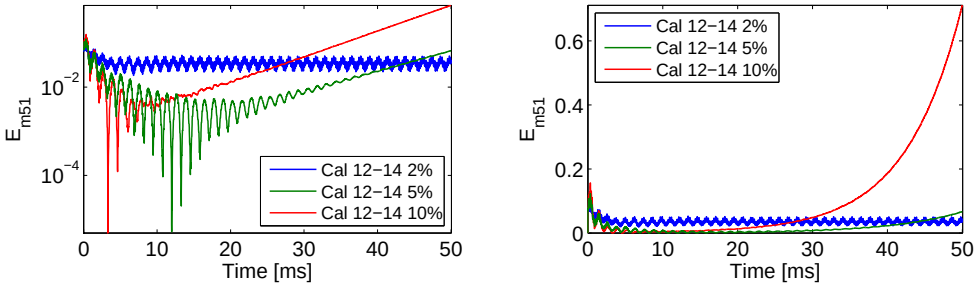


Figure 5.60: Effect of damping caliper mode 12 to 14 on the limit cycle obtained at 12 Bar, modal mechanical energies associated to mode C51

This chapter showed that damping 3 caliper modes has an effect similar to damping modes R44 and R45, and has the potential to tame the instability observed in chapter 4. Although the main instability can be avoided through this mean, other unstable modes can arise and generate other issues. The modification may nevertheless not be void, as the vibration levels of the new instability can be lower than the previous one. The damping ratio could also be optimized to obtain a regime where mode C44 becomes stable while mode C51 is less altered than in the simulations presented here.

Application of structural modification methods is a clear perspective to the current work. Relevant modification should be applied to the caliper to assess the level of frequency modification that can be attained. A realistic damping treatment of the caliper should also be performed to assess the damping ratios that could be obtained by such mean.

Damping application does not always provide a better stability in the case of brake squeal. This has been discussed in the literature and came out as a paradox for Sinou *et al.* [169, 170] which could be verified experimentally by Massi *et al.* [100]. For a complex structure like a full automotive brake assembly, the notion of proportional damping, discussed by the previous authors, can be seen in another light. Indeed, this notion involved keeping damping ratios constant between all DOF for a given modification. For the application presented in this chapter, some refinement can be added, using a comprehensive component-wise damping approach.

The risk when not applying proportional damping (which is impossible to realize in practice) is then to destabilize the system by several means. Destabilization can be understood as an increase of the limit cycle amplitude, although the modes damping ratio can be decreased. It is first possible to let another unstable mode express itself, with effects to assess. This pattern has been verified in practice and constitutes the most natural form of destabilization. In the same view, it is possible to alter the shape and damping ratios of other unstable modes, which can become predominant afterwards.

A second destabilization pattern has also been verified, *e.g.* looking at mode R45 damping. In such case, the unstable pattern remains the same, but the energy distribution between modes (or also components) constituting the instability is altered. This could also yield limit cycles of greater amplitude after a damping treatment.

6

Conclusion and perspectives

For the application of friction induced instabilities in automotive brakes, the thesis explored non-linear transient analyses and design oriented model reduction. Time simulation was made possible through two main components. First, novel reduction techniques were employed allowing the generation of very compact reduced systems. The reduction uses exact system modes at a nominal point, which leads to exact nominal predictions, and provides reduced interface models, which removes a common limitation of traditional CMS strategies.

The second contribution is linked to the time integration scheme. A non-linear Newmark scheme was modified through the introduction of a fixed Jacobian to treat contact behavior. The convergence properties have been shown to be very satisfying, in particular by avoiding the bouncing effect commonly encountered in Newmark schemes.

Classical models for contact-friction, based on a functional representation of contact pressure, were motivated by Bosch experimental knowledge in the domain and thus used here. A contribution of the thesis was an extension of the (1-ϵ)-compatibility framework used to handle contact in the presence of incompatible meshes. While the approach is not central to the thesis, it may become important when other non-linearities than the brake/pad contact will be considered.

With these developments, the stakes of the thesis are thus not to produce transient simulations, but to be able to exploit them. Various tools, namely modal sensors, modal damping, and space/time decomposition, were thus introduced to analyze time simulations. They allowed an insightful vision of modal interactions occurring in limit cycles, and illustrated the difference between modal growth rate and limit cycle amplitude.

The first results obtained by such analysis highlighted the actual links between complex modes and the limit cycle response. Specific studies developed the notion of modal competition, explaining why a single unstable mode can be observed in certain non-linear transient simulations while several complex modes are unstable. It was found however that all unstable modes could show a potential growth in the response. The mode with the highest growth rate tends to respond first and overtake other modes response.

Damping studies for non-linear systems has been under the spotlight for a few years. The notion of proportional damping [170] was in particular introduced by Sinou and Jezequel to separate cases where damping application is sure to have a positive effect and cases where non-linear studies were needed to assess the effect. Sometimes, destabilization occurred. This was referred to a paradox, where adding damping to the system had a negative effect on the limit cycle.

Proportional damping is too restricted to be applicable to complex structures. The only case where damping non-linear systems is sure to be positive is in fact when damping is proportional to stiffness. A contribution of the thesis was thus to introduce tools allowing specification of modal damping ratio for system or component modes with distinct values for each mode.

Stability results, obtained by linearizing the system at a specific nominal point, give complex modeshapes and poles. Poles with negative damping values are unstable and associated with a growth rate. The common practice is to optimize brake design to reduce or remove unstable poles of the system linearized around its static state. This habit was shown to be inappropriate in chapter 5 where increasing damping of mode R45 reduced the solution growth rate but increased the limit cycle amplitude. The need to use time simulations to evaluate the impact of damping changes is thus an important conclusion of this work and the proposed illustrations provide initial proposals for how to proceed.

The modal sensors and space/time decomposition for the limit cycle oscillations were shown to be well correlated and pointed specific modal participations to the non-linear responses. The results presented a limit cycle of dimension 2 in space, and only one unstable participation for the nominal state. These results are still coarse and a refinement was obtained by performing stability analysis at various instant in the limit cycle.

Since the contact-friction states evolve through time, so do complex modes. Transient stability diagrams showed that coalescence happens during the cycles, and that the growing unstable mode passed through stable phases during the cycle. A separation pattern was observed, which decreased the modal frequency and limited the effect of friction. The system mechanical energy was thus not constant in the limit cycle, but showed oscillations due to the mode coalescence/separation, with growth phases when the mode was unstable and decay phases when the mode was stable. The existence of a saturation in the amplitudes is thus attributed to these transitions and that seems an important conclusion.

The coalescence/separation pattern opens the discussion of modal interaction in the limit cycle. It was shown in particular that two real modes worked together to maintain the instability, although only one was energetic in the nominal response. A component-wise strain energy study of the transient limit cycle shape allowed refining this observation. A three component interaction, disc/pad/caliper, could be decomposed as a sequential disc/pad, then pad/caliper interaction. The most energetic interaction (pad/disc) was not the most critical interaction regarding the instability (pad/caliper). The importance of the whole structure in the expression of the contact-friction instability thus proved to be critical. This result justifies *a posteriori* the choice of keeping a complete brake assembly for simulations.

Since the cost of transient simulations remains important (around 12h), approximate non-linear simulations of minimal computation cost were proposed. The idea is to extrapolate the system behavior from trajectories associated with unstable complex modes. Although the corresponding non-linear efforts are computed, the mechanical equilibrium is not verified in these pseudo-cycles. The approach is clearly related to existing literature, for example the CNLMA [124] developed by Sinou, or the non-linear normal modes currently developed by Kerschen *et al.* [125].

Unstable mode ranking by evaluation the work of the non-linear forces over pseudo-trajectories gave interesting results. Pseudo-instant stability diagrams showed similar results to the actual transient simulations. Such study thus seems a promising low cost method to obtain an approximation of the limit cycles amplitudes each unstable mode could generate. The lack of mechanical equilibrium however limits the method relevance, and no saturation pattern could be observed for certain modes. Performing local mechanical equilibrium in the vicinity of the non-linear areas, when computing pseudo-cycles should resolve these problems and constitutes a clear perspective of this work.

While the need for time domain simulations has been well demonstrated, the relation between system modes and limits cycles still remains. Due to the links existing between system modes and the limit cycles, working in the frequency domain is relevant for design phases. Transient simulations are however mandatory to understand which modal behavior is critical for a nominal

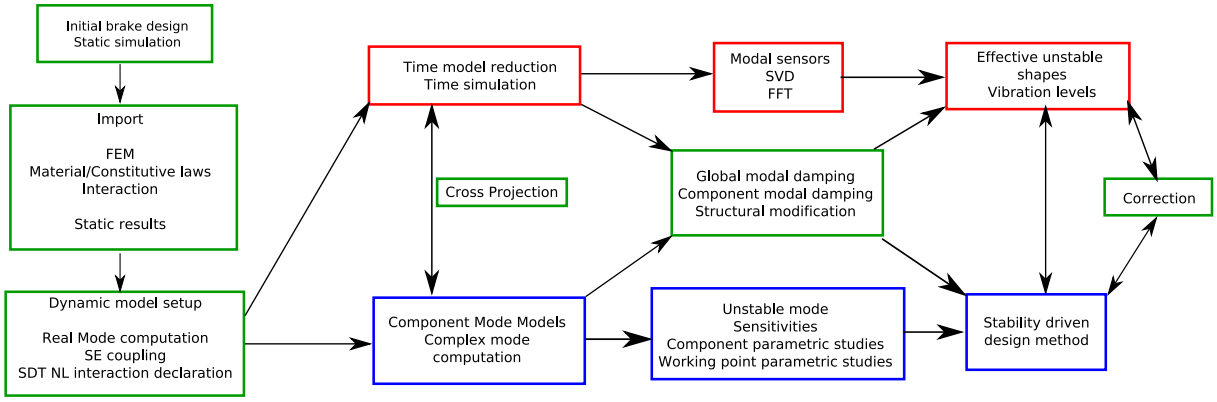


Figure 6.1: A suggestion of brake squeal simulation integrated in an NVH procedure

state, and to validate the redesign effect.

Considering the system components separately is a necessity at early design stages. Keeping such information throughout the V-cycle of design is thus critical. The CMT method, in which free/free component modes are explicit degrees of freedom of the reduced assembly, was introduced for this purpose. The historical methods by Rubin [50] or MacNeal [49] were not accurate. Realizing that assembled system modes are available, in particular with the development of multi-level eigensolvers [28], the method achieves exact modes while using component modes as DOF.

The reduction method thus allowed the generation of very compact models (around 1,300 DOF) with explicit local component-wise DOF. Parametric stability computations are therefore optimized and the effect of component redesign can be assessed with no need for full model evaluations. The approach is very useful to orient design studies. An extension illustrated in the thesis is the analysis of damping contributions from component modes in the time and frequency domains.

Figure 6.1 gives a synthesis of the tools developed throughout the thesis. Their combination provides an integrated NVH process to improve control of brake squeal noise problems at the design level. In this process, time and frequency simulations can be performed in parallel, using potentially different reduction strategies in both domains. Cross projection exists between both frameworks, allowing the correlation between time and frequency results.

From a new brake model, the stability can be assessed and compared to experimental data for potential updating, using the CMT. The stability analyses of the updated models can then be exploited in terms of sensitivity to give the most relevant design directions. A nominal time simulation is performed to evaluate modes predominant in the modal competition. The critical modal interactions and the limit cycle amplitude are thus obtained.

The effect of component design modification is then predicted in the frequency domain and validated by transient simulations, for which limit cycle continuations can be used to optimize the computations. Modal competition can thus be reassessed and new limit cycle amplitudes be computed. Physical modifications can eventually be performed and a full model simulation closes the proposed validation loop.

The integration of this process within the normal design tools used by Bosch is a clear perspective of this work. The CMT, corresponding to the blue branch of the process diagram of figure 6.1, has already been validated and deployed during the thesis. The time simulations detailed are in good agreement with Bosch experimental results. The validation of the red branch is thus partially obtained, while the simulation relevance has been demonstrated.

More precise experimental validation is another clear perspective of this work. Although interface characterization is impossible, component vibration levels, through Operational De-

flection Shape measurements should be validated. It is expected that this correlation will lead to refinements of the physical assumptions made.

The linearization of the brake model outside the pad/disc interface should sometimes be challenged. In particular in the limit cycles shown, amplitudes at other interfaces are such that assuming a linearized sticking contact is clearly not appropriate.

Implementing non-linearities at other interfaces is not a specific implementation challenge. The stakes are to quantify the level of refinement needed and assess the numerical cost. Implementation of non-linear forces at reduced interfaces can generate unwanted effects. The local effect is indeed lost as the generalized degrees of freedom will be distributed. If keeping all non-linear DOF does not appear possible, a perspective is the use of (1-ϵ) compatibility concepts to optimize the reduced behavior.

The implementation of viscoelastic materials in the time domain remains a challenge that may be needed in future work. First of all, it is common practice to use rubber shims in brakes [17] and their effect on the transient behavior will need to be assessed. Then, the lining material has very specific properties, and tends to have a viscoelastic behavior for automotive brake applications. Since it is a critical material, a better physical representation is probably needed.

A more systematic analysis of operational conditions is another clear need. Thermal effects, pad wear, hygrometry are for example currently neglected. The system is then studied in the very restrictive case of ambient temperature with unaltered components. Wear in particular is unavoidable, and can change the pad braking properties, geometry and positioning under static load, thus leading to significantly different component interactions.

Robust design will thus imply to address the need to build silent new brakes that maintain a controlled level of noise generation throughout their lifetime. Trade-offs will obviously depend on commercial strategies. With time simulations and design methodologies, the tools proposed in this work clearly contribute to improving the ability to base design decisions on quantitative predictions.

A

A design framework, complex modes and sensitivities

A.1 Complex modes computation

A mechanical assembly can be characterized by its modes. For stability evaluation, the computation of complex modes is necessary. In the frequency domain, system stability can be evaluated in the sense of Lyapunov (section 3.3) by looking at the sign of the real part of the complex eigenvalues. They are solution of the second order eigenvalue problem

$$[M]\{\psi_j\}\lambda_j^2 + [C]\{\psi_j\}\lambda_j + [K]\{\psi_j\} = \{0\} \tag{A.1}$$

where $[M], [C], [K], [\psi] = [\psi_j], [\Lambda] = diag(\lambda_j)$ are respectively the mass, damping, stiffness matrices, an eigenvector and an eigenvalue. Numerical libraries such as LAPACK [171] only support first order eigenvalue computations. One thus transforms (A.1) into a first order form. A classical possibility [157] is to use

$$\begin{bmatrix} C & M \\ M & 0 \end{bmatrix} \begin{bmatrix} \psi \\ \psi\Lambda \end{bmatrix} [\Lambda] + \begin{bmatrix} K & 0 \\ 0 & -M \end{bmatrix} \begin{bmatrix} \psi \\ \psi\Lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{A.2}$$

The definition of the orthogonality conditions is not as straight forward than for real modes. In particular, a potential lack of symmetry in the matrices (as sliding friction coupling, seen in section 3.2.4) yields different left ψ_l and right ψ_r complex modes depending on solving equation (A.1) or its transposed. Typically, the orthogonality conditions are written

$$\begin{cases} \psi_l^T C \psi_r + \Lambda \psi_l^T M \psi_r + \psi_l^T M \psi_r \Lambda = I \\ \psi_l^T K \psi_r - \Lambda \psi_l^T M \psi_r \Lambda = -\Lambda \end{cases} \tag{A.3}$$

The eigenmodes are then ψ_r, the frequencies $\omega_j = |\lambda_j|$, the damping and the real part of Λ are linked as $\Re(\lambda_j) = -\zeta_j|\lambda_j|$. A mode is unstable when its damping ratio is negative (positive real part). This is sometimes referred as a growth rate.

To lighten the complex mode computation, (A.1) is commonly reduced in the basis of the real modes of the associated conservative system. This imposes to know the symmetric stiffness matrix associated to the conservative system. However it is not always sufficient for systems showing strong antisymmetric coupling. Enhanced reduction methods are then necessary. Such method has been developed by Balmes [157, 41]. They are based on the Newton iteration concept.

Let $[K_S]$ and $[K_U]$ respectively be the symmetric (all symmetric coupling terms possible) and the unsymmetric part, and let $[\phi_j]$ be the result of

$$\left([K_S] - \omega_j^2 [M]\right)\phi_j = 0 \tag{A.4}$$

One can compute the modes ψ_{jR} within the reduction basis containing the modes $T = [\phi_j]$. One then has $\psi_j \approx [T]\{\psi_{jR}\}$ (subspace methods). The error associated with this reduction can be quantified in terms of residual force by inputting the modes $[T]\{\psi_{jR}\}$ in (A.1), yielding

$$\{R_{Fj}\} = \lambda_j^2 \{T\psi_{jR}\} + [C]\{T\psi_{jR}\}\lambda_j + [K_S]\{T\psi_{jR}\} + [K_U]\{T\psi_{jR}\} \tag{A.5}$$

At the first iteration, one can consider $\psi_j = \phi_j$ and $\lambda_j = i\omega_j$, and using (A.4), one finds a load residue given by

$$\{R_{Fj}\} = [C]\{\phi_j\} i\omega_j + [K_U]\{\phi_j\} \tag{A.6}$$

A displacement residue can be computed from (A.6). The idea, as for a Newton scheme, is to use an operator which allows to append the residue to the solution, $[K_S]$ is therefore chosen to obtain

$$\{R_{Dj}\} = [K_S]^{-1}\{R_{Fj}\} \tag{A.7}$$

If the displacement residue is not negligible it is appended to the reduction basis combined with an orthonormalization procedure,

$$[T_R] = [[\phi_j] \quad [R_{Dj}]]_{Orth} \tag{A.8}$$

In many cases, (A.1) can be simplified, which allows optimized solutions to be used. In a case without viscous damping, (A.1) is already a first order eigenvalue problem, with square eigenvalues

$$[M]\{\psi_j\}\lambda_j^2 + [K]\{\psi_j\} = \{0\} \tag{A.9}$$

Further simplification can be performed when the mass matrix is the identity. In that case, (A.9) consists in the eigenvalue extraction of $[K]$.

A.2 Sensitivity computation

The goal of design is to evaluate the variation trends of the nominal complex modes as function of parameters p. A basic idea to obtain quick trend evaluations is to compute the mode sensitivity to p by deriving equation (A.1). The sensitivity is expressed as $\frac{\partial\lambda_j(p)}{\partial p}$. The variation of the mode shapes is not studied but only the effect on the eigenvalues.

The derivation of equation (A.1) yields

$$\left[\frac{\partial[M]}{\partial p}\lambda_j^2 + \frac{\partial[C]}{\partial p}\lambda_j + \frac{\partial[K]}{\partial p} + 2[M]\lambda_j\frac{\partial\lambda_j}{\partial p} + [C]\frac{\partial\lambda_j}{\partial p}\right]\{\psi_{rj}\} = [Z(\lambda_j)]\frac{\partial\{\psi_{rj}\}}{\partial p} \tag{A.10}$$

where $[Z] = [M]\lambda_j^2 + [C]\lambda_j + [K]$. Equation (A.10) is of the type $Zq = B$ and admits solutions if and only if B is orthogonal to the kernel of Z^T, *i.e.* if $B^T Z = 0$. The left hand side of (A.10) multiplied by the corresponding left complex modes ψ_{lj}^T for any j yields

$$\left\{\psi_{lj}^T\right\}(2[M]\lambda_j + [C])\{\psi_{rj}\}\frac{\partial\lambda_j}{\partial p} = \left\{\psi_{lj}^T\right\}\left[\frac{\partial[M]}{\partial p}\lambda_j^2 + \frac{\partial[C]}{\partial p}\lambda_j + \frac{\partial[K]}{\partial p}\right]\{\psi_{rj}\} \tag{A.11}$$

Using the first orthogonality condition in (A.3),

$$\left\{\psi_{lj}^T\right\}(2[M]\lambda_j + [C])\{\psi_{rj}\} = 1 \tag{A.12}$$

which yields the formulation of the complex modal sensitivity

$$\frac{\partial\lambda_j}{\partial p} = \left\{\psi_{lj}^T\right\}\left[\frac{\partial[M]}{\partial p}\lambda_j^2 + \frac{\partial[C]}{\partial p}\lambda_j + \frac{\partial[K]}{\partial p}\right]\{\psi_{rj}\} \tag{A.13}$$

Equation (A.13) clearly gives the eigenvalue variation as function of the system matrices variations, for a parameter p. For instability studies, it is trivial from this expression to evaluate the instability sensitivity by directly exploiting the real part of $\frac{\partial\lambda_j}{\partial p}$.

B

Iterative methods for real mode computation

The eigenvalue problem is well known in the field of structural dynamics, as it provides the real modes of a structure. Such resolution is the necessary first step to most applications, for FRF computation, system qualification, model reduction, stability computation, *etc.*

A linear system is characterized by its stiffness K and mass M matrices. Both are symmetric definite positive. The system modes correspond to the space decomposition of the structure free response, and verify

$$(K - \omega_j^2 M)\phi_j = 0 \tag{B.1}$$

where couples (ω_j, ϕ_j) are unknown.

The problem (B.1) is a very classical problem in mathematics. The resolution target here very large sparse matrices, which is the common topology resulting from the finite element discretization. Full resolutions of system (B.1) are therefore too computationally demanding for full solvers, like the QR solver.

Partial solvers have been developed to tackle this issue. The objective is then to obtain an approximation of the first modes of the structure. Partial solvers are iterative and have been extensively described in the literature, for example by Saad [51].

The inverse power method [139], notes that equation (B.1) can be recast as

$$K^{-1}M\phi_j = \frac{1}{\omega_j^2}\phi_j \tag{B.2}$$

The resulting inverse problem is then solved by constructing a sequence of solutions

$$\phi^{n+1} = K^{-1}M\phi^n \tag{B.3}$$

which converges towards the smallest system eigenvalue of the system. Once convergence is satisfying, the second smallest eigenvalue can be found by looking for a second solution of (B.2), orthogonal to the previous one.

Convergence problems appear quickly using such method, as maintaining the orthogonality of the whole family of eigenvectors is generally difficult. Besides, if some eigenvalues are close, the convergence properties in $\mathcal{O}(\frac{\omega_j^2}{\omega_{j+1}^2})$ can make the method underperforming.

Subspace methods have been developed for example by Bathe and Wilson [172]. The resolution method is based on the assumption that the n first modes ϕ_j of the structure can be approximated by a set of shapes $\tilde{\phi}_j$, linked to the solution by a matrix T, such that

$$\tilde{\phi}_j \approx T\phi_j \tag{B.4}$$

The method of Bathe thus iterates on the basis T by using the inverse power method (B.2), where the whole basis T is iterated instead of a single eigenvector.

The most widespread iterative solvers have been developed in the 50's and are based on enhancement methods, using the algorithms of Lanczos [139] or Arnoldi [173]. Starting from a random vector z_1, one generates the Krylov sequence

$$T^k = \left[z_1,\ z_2 = (K^{-1}M)z_1,\ z_3 = (K^{-1}M)^2 z_1, \cdots,\ z_{k+1} = (K^{-1}M)^k z_1\right] \tag{B.5}$$

which should converge towards the system eigenmodes. In practice, the vectors generated are very collinear and thus necessitates an orthogonalization at each step of the process. The method is then close to conjugate gradients methods employed for hermitian problems.

The methods of Arnoldi and Lanczos differ on the type of orthogonalization made, since a total orthogonalization of z_k against T^k is performed by the former, and only a mass orthogonalization is performed by the latter.

The Lanczos method is efficient enough so that most recent developments concerned its improvement. The IRA method developed by Sorensen [174] aims at avoiding a recurrent problem of the enhancement methods. Indeed, the iterative generation of the Krylov sequence (B.5) is not limited in size and tends to become too large. For a large number of iteration, numerical conditioning is lost and the orthogonalization process becomes difficult. The idea is then to introduce restart phases where the iterated basis is reset with a new starting vector based on the eigenmodes already found.

The residue iteration method developed by Balmes and Bobillot [41], is between subspace iterations and enhancement methods. Its application is presented in appendix A and is therefore not reproduced here in details. The method considers an enhancement of the mode basis but an orthogonalization process like the Arnoldi method. The residue iteration method is commonly employed in SDT to improve convergence of the Lanczos solver implemented. It is also used for complex mode computation, as presented in appendix A.

Gain in performance for eigenvalues resolution algorithm is now sought using domain decomposition techniques. This is in phase with the current computer developments preferring parallel computations over pure performance. The AMLS developed by Kaplan and Bennighof [28, 29] and presented in section 2.4 thus aims at reducing super-large systems to apply the Lanczos algorithm to systems of acceptable sizes.

C

Global parametric studies of brake systems

C.1 Friction parametric studies

This section presents the specific case of friction taken as a parameter. The unsymmetric friction coupling matrix presented in section 3.2.4 represents the tangential contact forces between the disc and the pads. It can be factorized by the friction coefficient. Consequently, the parametric study concerns the linear variation of the unsymmetric stiffness matrix. Results are plotted in figure C.1

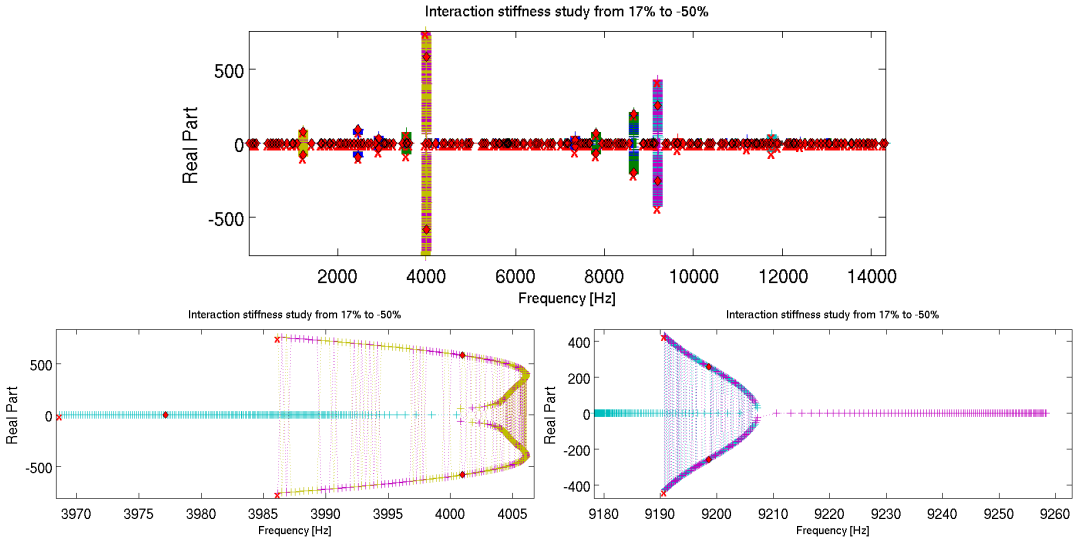


Figure C.1: Stability diagram as function of the friction coefficient, from 0.7 to 0.3. (x) for $\mu = 0.7$; ($\diamond$) for the nominal points ($\mu = 0.6$)

The friction coefficient is a classical parameter for brake systems stability analysis however; the global friction coefficient (computed using contact force resultants) may vary despite a constant local coefficient, as presented by Linck [8]. This is possible due to the brake vibration that generates a variation of the friction forces in the Coulomb cone for a constant normal contact state.

In our case, such parametric study is useful to highlight the unstable modes especially sensitive to the contact configuration. Indeed these modes could be triggered by the system vibrations. Figure C.1 spots 10 such modes, two of which are very unstable at 4 and 9.2 kHz. This is well correlated to the tests performed by Bosch (see section 1.2).

As expected, a higher friction coefficient generates unstable modes with higher real parts.

The frequencies remain close to their nominal values. Conversely, lower friction coefficient coefficients yield lower real parts, with for example a stiff decrease of the instability level of the 4 kHz mode under $\mu = 0.53$.

C.2 Pressure parametric studies

This section presents different strategies for parametric studies in pressure. The aim is to estimate the evolution of the stability diagram between fully computed design points using interpolation techniques. This allows efficient mode tracking to properly identify the unstable modes and their evolution over large pressure spans.

Two main features of the component mode reduced model (from section 2.3) are depending on the braking pressure, the reduction basis enhancement and the reduced interaction matrices. The latter is commonly handled by interpolating the matrices computed at different design points. The interpolation concept is presented in figure C.2, a piece-wise linear interpolation and a 3 point quadratic interpolation are tested in the following.

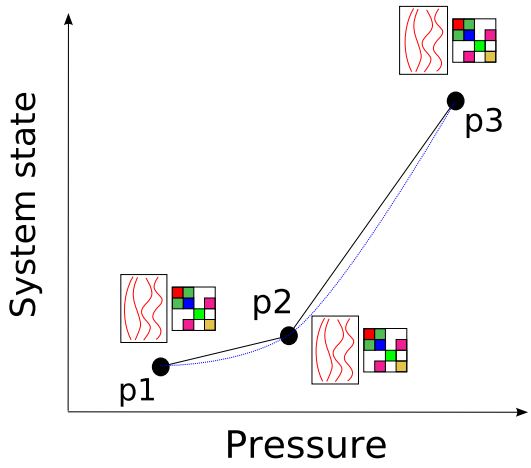


Figure C.2: Interpolation strategies of the system states as function of the pressure

The reduction basis enhancement issue is combined to the interpolation since all matrices must be reduced in the same basis. A basic solution would be to use the highest pressure point to reduce the model. However, it is not sufficient in many cases; a classical improvement is to use a multi-pressure basis. The choice of the design points to include is discussed in the following.

In practice, the reduction basis is computed by component. The multi-pressure enhancement is performed by appending one pressure at a time in decreasing order, with an orthonormalization procedure. All matrices are then reduced in the final basis.

The multi-pressure enhancement can yield much larger reduction basis, which is an issue in terms of computing performance. The reduction in the system real modes (as presented in appendix A) for the complex mode computation is then unavoidable. Further optimization is presented in the following to lighten the computation. For piecewise linear interpolation procedure, one can split the job by considering only two pressures at a time. For other strategies, residue iteration techniques such as presented in [41] would lighten the process by updating the second layer reduction basis instead of recomputing it. Such refinements are perspectives of the present work.

The results presented here are studying the effect of the pressure from 3 to 24 bars using design points at 24, 20, 16, 12, 8, and 3 bars. The computation has been split into 5 sub jobs, for each segment. The complex modes were computed using two reduction layers. A cutoff frequency was given at 17.5 kHz. The component reduction basis (first layer) was made of the

reduction basis at the higher pressure enhanced with the reduction basis of the lower pressure, for each component. Figure C.3 shows the main result.

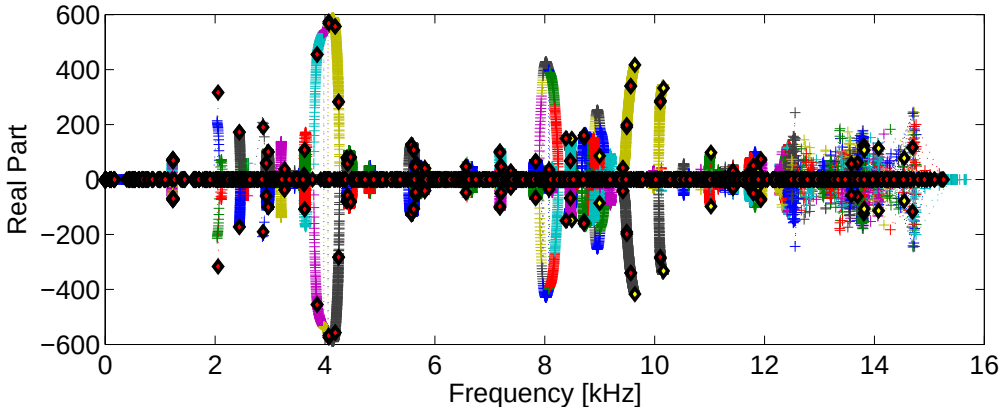
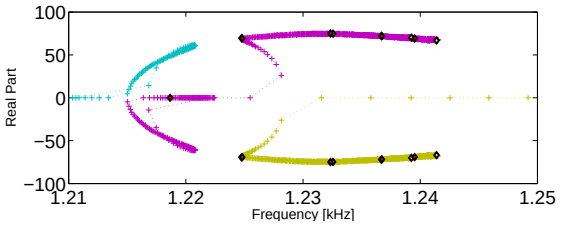
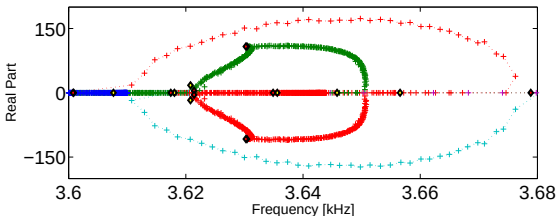


Figure C.3: Stability as function of the pressure, from 3 to 24 bars

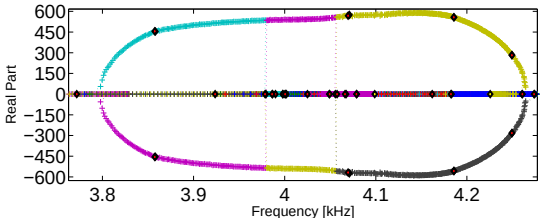
Patterns showing instability for specific pressure ranges are observed, the main instabilities are found in the 2-5 kHz and 8-10 kHz ranges. This allows mode tracking to ascertain the modes to alter, in CMT analyzes for example. Figure C.4 features zoomed in modes of the full diagram. The design points have been appended to highlight the pressure ranges transitions. The pole evolutions are continuous; a slight gap can appear between two sub-jobs.



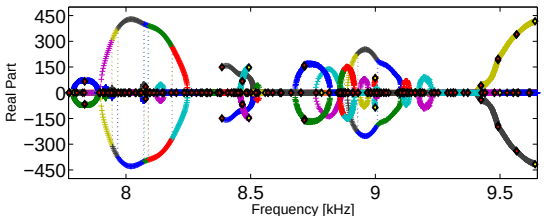
(a) Mode 19, 20, 21



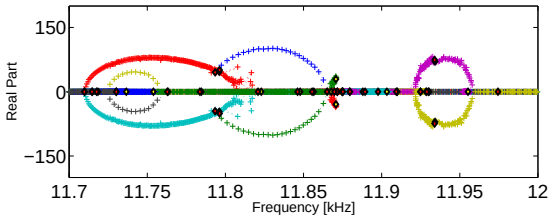
(b) Mode 34, 35



(c) Mode 50, 51, 52, 53



(d) Modes 106 to 132



(e) Modes 162 to 166

Figure C.4: Stability diagrams as function of the pressure, from 3 to 24 bars. Black edged yellow $\diamond$ (*resp.* black edged red $\diamond$) stand for the first (*resp.* last) design points of a sub-job

Although the curves featured are continuous, there are threshold effects particularly in figure C.4a or figure C.4b. These are due to the changes between to sub-jobs, which involves the reduction basis and the interaction matrix interpolation switch. Each sub-job is giving a relevant result, and the overall diagram gives many clues on the system's variation as function of the pressure however, the sub-job transitions will need to be addressed in the future.

Several interpolation strategies can be imagined, and some variants are tested in figure C.5. Between 8 and 12 Bar, a direct interpolation using linear or 3 points quadratic strategies are compared to piecewise applications.

The variation of the interpolation strategy behaves as expected, with the accuracy obtained, a compromise with the model size (computation times). The error observed with the simplest strategy is relatively small, however the larger the pressure interval the coarser the result. Besides, the conditioning can be worsened when using too many reduction bases.

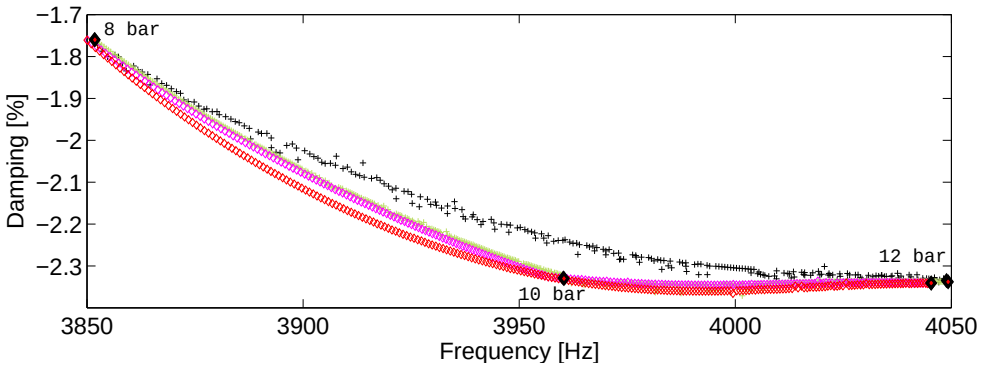


Figure C.5: Various interpolation strategies between 8 and 12 bar. + Linear interpolation between 8 and 12 bar. + Piecewise interpolation between 8 and 10 bars then 10 and 12 bar. ◇ Piecewise interpolation between 8 and 12 bar using 3 reduction bases at 8, 10 and 12 bar. ◇ 3 points quadratic interpolation using 8,10 and 12 bar

D

Macroscopic friction models

Friction phenomena at the scale of multibody dynamics generate various and complex effects, such as hysteresis, dwell, *etc.* This has led some authors to propose analytical friction models implementing all these phenomena. In the 70's, Dahl [175] was amongst the first to propose continuous models featuring hysteresis allowing better stick-slip transitions. Recent application exists in the dry friction domain, as Chevallier [176] used it to simulate clutch judder.

The experimental assessment of friction phenomena improved a lot in the last two decades, through the impulsion of mechatronics, which encouraged the generation of continuous friction models featuring diverse richness degrees. Indeed, the precision needed for robotics command requires taking actioners friction into account. Armstrong-Hélouvry in 1991 initiated the movement and provided a survey of friction models [177].

The control problem is rather complex. Continuous laws are needed to comply with automats programs, which must properly handle movement starts (static-dynamic transitions), response times with oscillations due to friction and so on. The great advantage of mechatronics applications come from the experiments. Indeed, friction behavior is studied as a global phenomenon for each interface. Its effect is analyzed through time responses from actioner engine commands and displacement-velocity responses of the component actioned. The PhD of Borsotto [178] provides a good insight of the state-of-the-art in the domain of mechatronics control taking friction into account.

The experimental characterization allowed classifying whose properties can be expected from a contact model. Seven specific types of behavior can be distinguished, namely

- **Stick** – ability to represent friction forces for sticking states
- **Rising static friction** – ability to represent higher static friction coefficients (also named *dwell*)
- **Elastic friction** – ability to represent stick-slip transitions
- **Non drifting** – ability to represent a constant position for stick states
- **Stribeck effect** – ability to represent the friction force loss at the beginning of sliding
- **Frictional memory** – or friction lag, ability to represent the hysteresis in stick-slip transitions

- **Viscous friction** – ability to represent friction forces for sliding states with lubrication

The friction models presented below can be distinguished between static and dynamic models. Dynamic models are of course the finality, but the static models are of interest as they illustrate specific effects more simply than the dynamic ones. Of course the Coulomb model, equation (3.2) is the simplest one, from which enhancements are sought.

Static models are not trying to give a simple expression of the friction force at zero sliding velocity. As for the Coulomb model, in the case where the point is sticked, the friction force is computed as the reaction to other mechanical forces, so that it is not determined by the law in practice.

First of all the rising static friction effect can be implemented by assuming a greater friction coefficient μ_s in the case of sticking, the Coulomb with static friction is thus written

$$\begin{cases} \|f_T\| < \mu_s \|f_N\| \Leftrightarrow w = 0 \\ \|f_T\| = \mu \|f_N\| \Leftrightarrow f_T = \mu \|f_N\| \frac{w}{\|w\|} \end{cases} \tag{D.1}$$

where $\mu_s \geq \mu$. The Coulomb + static friction is plotted in figure D.1. The viscous friction effect can then be added, using a parameter χ, representing a linear increase of friction forces with the velocity, writing

$$\begin{cases} \|f_T\| < \mu_s \|f_N\| \Leftrightarrow w = 0 \\ \|f_T\| = \mu \|f_N\| + \chi \|w\| \Leftrightarrow w \neq \{0\} \end{cases} \tag{D.2}$$

which is also plotted in figure D.1.

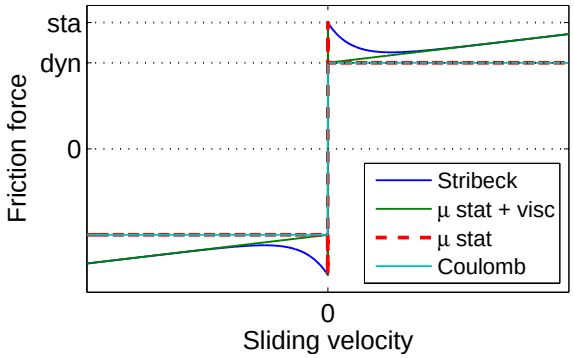


Figure D.1: Illustration of friction static models

The continuity problems due to the static/dynamic friction coefficients can eventually be assessed, through the Stribeck effect, which proposes a continuous transition between both systems,

$$\begin{cases} \|f_T\| < \mu_s \|f_N\| \Leftrightarrow w = 0 \\ \|f_T\| = \left(\mu + (\mu_s - \mu) e^{-\|\frac{w}{v_s}\|^\delta}\right) \|f_N\| \text{sign}(w) + \chi \|w\| \Leftrightarrow w \neq 0 \end{cases} \tag{D.3}$$

using fitting parameters v_s and δ.

From the considerations of the static systems, the numerical issue of detecting a null sliding velocity can be raised. Since all the static friction models presented rely on two definitions depending on the friction state, the detection must be robust, which can pose numerical problems. Karnopp [179] proposes to introduce a tolerance on the detection of the zero sliding velocity, using δv, so that inside the interval $[-\delta v \ ; \ \delta v]$ a sticking state is declared, other effects can be setup as usual. Using static friction and linear viscous forces, one has

$$\begin{cases} \|f_T\| < \mu_s \|f_N\| \Leftrightarrow \|w\| \leq \delta v \\ \|f_T\| = \mu \|f_N\| + \chi \|w\| \Leftrightarrow \|w\| > \delta v \end{cases} \tag{D.4}$$

The static models still feature a discontinuity at stick-slip transitions and lack friction lag. A class of dynamic models however exist, which must take into account the pre-sliding movements and hysteresis for stick-slip.

The general formulation of such system is given for example by Al-Bender *et al.* [180], relating the heuristics of friction models. The use of internal states is necessary, so that 2 equations describe a friction force; the friction equation itself

$$f_T = \mathcal{F}\left(\{z\}, w, q\right) \tag{D.5}$$

and the internal states equation

$$\frac{d\{z\}}{dt} = \mathcal{G}\left(\{z\}, w, q\right) \tag{D.6}$$

where $\{z\}$ is a vector of internal states, w the sliding velocity, and q the displacement vector.

Equations (D.5) and (D.6) represent the class of so-called rate-state models, and originated in the 80's for geophysics applications, in particular by Rice and Ruina [181]. Heuristics give some indication on the expression of $\mathcal{F}$ and $\mathcal{G}$. In the case of constant velocity, $\mathcal{F}$ is only function of the sliding velocity, as the state vector remains constant. In the case of sticking, $\mathcal{F}$ is a hysteretic function of the displacement vector.

The pre-sliding effect was first introduced by Dahl [175], who provided a model inspired by plasticity deformation formulations. The friction force is then described by

$$\frac{df_T}{dt} = \frac{df_T}{dz}\frac{dz}{dt} \tag{D.7}$$

where z is the tangential relative displacement whose derivative $w = \frac{dz}{dt}$ is the sliding velocity. The expression of the friction force is a slope function, written

$$\frac{df_T}{dx} = \sigma\left(1 - \frac{f_T}{f_g} sign(w)\right)^{\alpha} \tag{D.8}$$

where $f_g = \mu\|f_N\|$ is the sliding state friction force. The slope at the origin σ, and a material parameter α are fitting parameters. α commonly depends on the material. Relation (D.7) is then explicitly written as

$$\frac{df_T}{dt} = \sigma\left(1 - \frac{f_T}{f_g} sign(w)\right)^{\alpha} w \tag{D.9}$$

The contact force is an unknown to the system as for the Lagrangian formulation but resolved using an ordinary differential equation. The hysteretic behavior of the Dahl law is illustrated in figure D.2 where the evolution of f_T is plotted for an harmonic excitation of the displacement z.

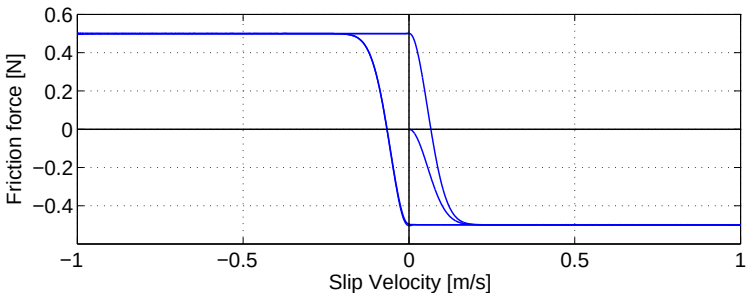


Figure D.2: Hysteretic effect of the Dahl model

From this first model, several aspects are missing, such as the non-drifting, rising static friction, Stribeck, and viscous friction. A generalization was proposed by Canuda de Wit *et al.* [182] presenting the LuGre model.

The friction force equation is enhanced to feature a damping term and the viscous forces, explicitly using an internal state z, writing

$$f_T = \sigma_0 z + \sigma_1 \dot{z} + \sigma_v w \tag{D.10}$$

where σ_0 and σ_1 represents a stiffness and damping term relative to the internal state, and σ_v is the viscous force coefficient. The internal states are defined by

$$\frac{dz}{dt} = \left(1 - \frac{\sigma_0 z}{g(w)} sign(w)\right) w \tag{D.11}$$

the function g representing the sliding regime (the role of f_g in the Dahl model). Taking the Stribeck effect into account,

$$g(w) = \left(\mu + (\mu_s - \mu) e^{-\| \frac{w}{w_s} \|}\right) \tag{D.12}$$

with μ and μ_s respectively dynamic and static Coulomb coefficient, w_s a fitting parameter.

The LuGre model, described by equations (D.10), (D.11) and (D.12) and shown in figure D.3 provides a better representation of experimental observations, which were possible by the installation of new tribometers, as presented by Lampaert *et al.* [183].

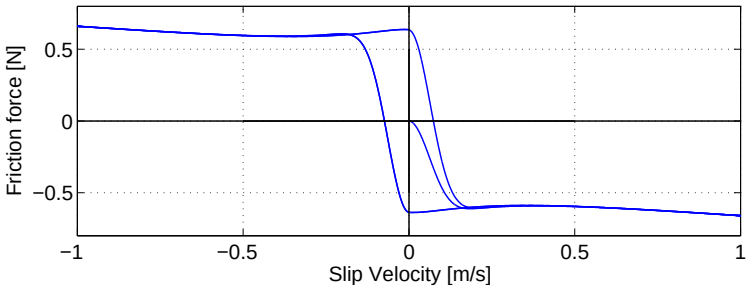


Figure D.3: Behavior of the LuGre model

The LuGre model is not fully satisfying as it lacks the non-drifting property. Enhancements were then proposed by several authors, as the Leuven model by Swevers *et al.* [184] deemed difficult to implement. Bliman and Sorine [185] proposed another generalization of the Dahl model, allowing it to gain an order. Another formalism was introduced by Haessig and Friedland [186] who proposed the Bristle model and the Reset Integrator model. The Bristle model is based on the modeling of contact as a simplified set of springs whose link can break with the tangential load. The latter model is a pure behavior model based on automatics and control formulations. An elastoplastic formulation was eventually proposed by Dupont [187]. These models all feature a specific modeling of the friction force global behavior and generally provide a subset of the expected properties presented. They can classically be expressed using the general expressions given in equations (D.5) and (D.6).

A new class of advanced friction models is based on Maxwell cells. Their advantage relies in the simplicity of their implementation and their fidelity to the friction behavior. The model, called *Generalized Maxwell-Slip* has been presented by Al-Bender, Lampaert and Swevers [180]. Such model is of clear interest and is the natural improvement direction sought for the present thesis perspectives.

The model, illustrated in figure D.4 proposes to consider a set of elasto-plastic cells, each cell having a saturation threshold g_i, spring stiffness k_i and internal state z_i giving the global friction force representation of figure D.4.

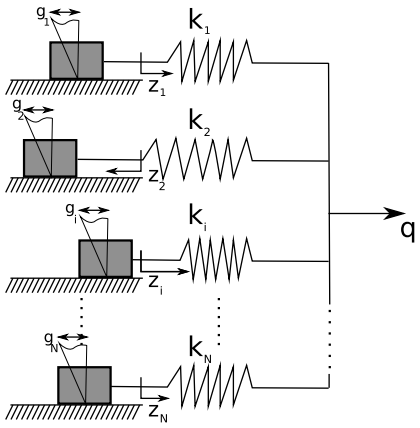


Figure D.4: Illustration of the Maxwell-Slip model using N elasto-plastic cells

The Generalized Maxwell-Slip model is formulated as

$$f_T(t) = \sum_{i=1}^{N} (k_i z_i + c_i \dot{z}_i) + f(w) \tag{D.13}$$

where k_i and c_i are a stiffness and damping term associated to the spring of each cell, and $f(w)$ represents the viscous force, generally taken constant or linear.

The internal state is defined as function of the stick/slip state of each cell, writing

$$\begin{cases} \frac{dz_i}{dt} = w \quad \text{if sticking and} \quad \|z_i\| \leq g(w) \\ \\ \frac{dz_i}{dt} = \left(1 - \frac{z_i}{g(w)}\right) A_i sign(w) \quad \text{if sliding and until} \quad w = 0 \end{cases} \tag{D.14}$$

A_i is then a fitting coefficient, and $g(w)$ represents the Stribeck effect as defined in equation (D.12).

The formulation can then be seen as a coupling of LuGre models – with sticking handling – itself a generalization of the Dahl model. The non-drifting property is well handled by the sliding/sticking conditions of each internal state. All features expected from a friction model are available.

Figure D.5 plots a sample response of the model, using 4 cells. The drawback of such model is the number of parameters to identify, as each cell shows 3 to 4 parameters to fit to potential data. Simplifications can be setup for the parameters, by implementing a base set of parameters for all cells and a single scaling coefficient for each.

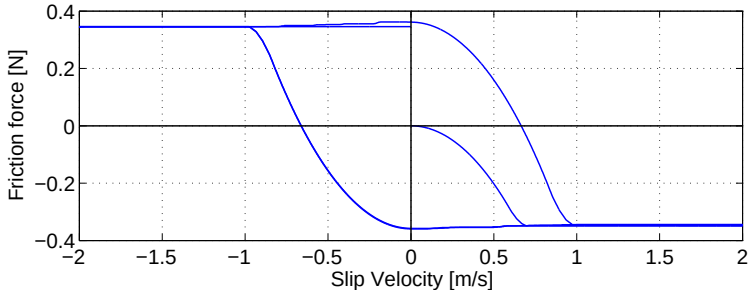


Figure D.5: Sample behavior of the Generalized Maxwell-Slip model

The time integration of models with internal states and first order equations is not trivial regarding classical implicit time integration schemes, as presented in chapter 4. The difficulties raised should first be solved before using such models in brake squeal applications, but their implementation seems promising.

Bibliography

[1] A. Akay, "Acoustics of friction," *The Journal of the Acoustical Society of America*, vol. 111, no. 4, pp. 1525–1548, 2002.

[2] K. Popp, N. Hinrichs, and M. Oestreich, "Analysis of a self excited friction oscillator with external excitation," *Dynamics with friction: Modeling analysis and experiment*, vol. 7, pp. 1–35, 1996.

[3] M. North, "Disc brake squeal," *Proceeding of IMechE*, pp. 169–176, 1976.

[4] D. Vola, *Frottements et instabilités en dynamique : bruits de crissement*. PhD thesis, Université d'Aix-Marseille II, 1998.

[5] F. Moirot, *Etude de la stabilité d'un équilibre en présence de frottement de Coulomb*. PhD thesis, Ecole Polytechnique, 1998.

[6] X. Lorang, *Instabilité vibratoire des structures en contact frottant: Application au crissement des freins de TGV*. PhD thesis, Ecole Polytechnique, 2007.

[7] F. Massi, *Dynamic and tribological analysis of brake squeal*. PhD thesis, INSA Lyon, 2006.

[8] V. Linck, *Modélisation numérique temporelle d'un contact frottant*. PhD thesis, INSA Lyon, 2005.

[9] A. Meziane, *Apport des analyses numériques temporelle et fréquentielle dans l'étude des instabilités de contact - Validation expériementale*. PhD thesis, INSA Lyon, 2007.

[10] G. Fritz, *Etude des phénomènes de crissement pour les freins automobiles - Modélisation non-linéaire et conception robuste*. PhD thesis, Ecole Centrale de Lyon, 2007.

[11] N. Coudeyras, *Analyse non-linéaire des instabilités multiples aux interfaces frottantes : application au crissement de frein*. PhD thesis, Ecole Centrale Lyon, 2009.

[12] H. Ouyang, J. Mottershead, D. Brookfield, and M. James, S.and Cartmell, "A methodology for the determination of dynamic instabilities in a car disc brake," *International Journal of Vehicle Design*, vol. 23, no. 3-4, pp. 241–262, 2003.

[13] H. Ouyang, A. Abu-Bakar, and L. Lijie, "A combined analysis of heat conduction, contact pressure and transient vibration of a disk brake," *International Journal of Vehicle Design*, vol. 51, no. 1-2, pp. 190–206, 2009.

[14] N. M. Kinkaid, O. M. O'Reilly, and P. Papadopoulos, "Automotive disc brake squeal," *Journal of Sound and Vibration*, vol. 267, no. 1, pp. 105–166, 2003.

[15] J. Fieldhouse and W. a. T. C. Steel, "The measurement of the absolute displacement of a noisy disc brake," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 222, no. 7, pp. 1121–1140, 2008.

[16] N. Hoffmann and L. Gaul, "Effects of damping on mode-coupling instability in friction induced oscillations," *ZAMM - Journal of Applied Mathematics and Mechanics / Zeitschrift für Angewandte Mathematik und Mechanik*, vol. 83, pp. 524–534, 2003.

[17] H. Jun and T. Gang-Feng, "Complex mode analysis on disc brake squeal and design improvement," in *SAE paper, No. 2009-01-2101*, 2009.

[18] M. Tan, K. Stepper, K. Abdelhamid, A.-H. Afaneh, and P. Blaschke, "An integrated process for design against brake squeal," in *International Modal Analysis Conference (IMAC XXI)*, 2003.

[19] N. Roy, S. Germes, B. Lefbvre, and E. Balmes, "Damping allocation in automotive structures using reduced models.," *PDF*, September 2006.

[20] A. Sternchüss, *Multi-level parametric reduced models of rotating bladed disk assemblies.* PhD thesis, Ecole Centrale de Paris, 2009.

[21] G. Vermot des Roches and E. Balmes, "Time simulation of squeal phenomena in realistic brake models," in *Proceedings of the International Conference on Advanced Acoustics and Vibration Engineering (ISMA)*, pp. 3007–3019, 2008.

[22] K. Mahmoud, "An efficient approach to structural optimization," *Computers & Structures*, vol. 64, no. 1-4, pp. 97–112, 1997.

[23] E. Balmes, "Parametric families of reduced finite element models. theory and applications," *Mechanical Systems and Signal Processing*, vol. 10, no. 4, pp. 381–394, 1996.

[24] N. Maia and J. Silva, *Theoretical and Experimental Modal Analysis.* John Wiley & Sons, 1997.

[25] M. Corus, *Amélioration des méthodes de modification structurale par utilisation de techniques d'expansion et de réduction de modèle.* PhD thesis, École Centrale Parparis, 2003.

[26] B. Groult, *Extension d'une méthode de modification structurale pour la conception de dispositifs dissipatifs intégrant des matériaux viscoélastiques.* PhD thesis, École Centrale Paris 2008-14, 2008.

[27] G. Vermot des Roches, J.-P. Bianchi, E. Balmes, R. Lemaire, and T. Pasquet, "Using component modes in a system design process," in *International Modal Analysis Conference (IMAC XXVIII)*, 2010.

[28] M. Kaplan, *Implementation of Automated Multilevel Substructuring for Frequency Response Analysis of Structures.* PhD thesis, The University of Texas at Austin, 2001.

[29] J. Bennighof, M. Kaplan, M. Muller, and M. Kim, "Meeting the nvh computational challenge: Automated multi-level substructuring," *International Modal Analysis Conference*, pp. 909–915, 2000.

[30] G. Vermot des Roches, E. Balmes, H. Ben Dhia, and R. Lemaire, "Mesure de compatibilité et recherche de solutions régulières en contact pénalisé," in *Neuvième Colloque National en Calcul des Structures (Giens2009)*, vol. 2, pp. 683–688–3019, 2009.

[31] G. Vermot des Roches, E. Balmes, H. Ben Dhia, and R. Lemaire, "Compatibility measure and penalized contact resolution for incompatible interfaces," *European Journal Of Computational Mechanics*, vol. 19, pp. 317–329, 2010.

[32] G. Vermot des Roches, E. Balmes, R. Lemaire, and T. Pasquet, "Design oriented time/frequency analysis of contact-friction instabilities in application to automotive brake squeal," in *Proceedings of the International Conference on Advanced Acoustics and Vibration Engineering (ISMA)*, pp. 4383–4398, 2010.

[33] G. Vermot des Roches, E. Balmes, R. Lemaire, and T. Pasquet, "Design oriented simulation of contact-friction instabilities in application to realistic brake assemblies," in *Sixth European conference on braking (JEF)*, November 2010.

[34] H. Ben Dhia and G. Rateau, "The arlequin method as a flexible engineering design tool," *Int. J. Numer. Meth. Eng.*, vol. 62, pp. 1442–1462, December 2005.

[35] A. Combescure and A. Gravouil, "A time-space multi-scale algorithm for transient structural nonlinear problems," *Mécanique & Industries*, vol. 2, no. 1, pp. 43–55, 2001.

[36] L. Bureau, C. Caroli, and T. Baumberger, "Elasticity and onset of frictional dissipation at a non-sliding multicontact interface," in *Royal Society A: Mathematical, Physical and Engineering Sciences, 459 (2039)*, pp. 2787–2805, 2003.

[37] L. Bureau, T. Baumberger, and C. Caroli, "Non-amontons behavior of friction in single contacts," *Eur. Phys. J. E*, vol. 19, no. 2, pp. 163–169, 2006.

[38] G. Vermot des Roches, E. Balmes, R. Lemaire, and T. Pasquet, "Designed oriented time/frequency analysis of contact friction instabilities in application to automotive brake squeal," in *Vibrations, Chocs & Bruit (VCB XVIIth symposium)*, 2010.

[39] R. J. Craig, "A review of time-domain and frequency domain component mode synthesis methods," *Int. J. Anal. and Exp. Modal Analysis*, vol. 2, no. 2, pp. 59–72, 1987.

[40] D. de Klerk, D. Rixen, and S. N. Voormeeren, "General framework for dynamic substructuring : History, review and classification of techniques," *AIAA Journal*, vol. 46, no. 5, pp. 1169–1181, 2008.

[41] A. Bobillot, *Méthodes de réduction pour le recalage. Application au cas d'Ariane 5.* PhD thesis, École Centrale Paris, 2002.

[42] C. Florens, *Modeling of the viscoelastic honeycomb panel equipped with piezoelectric patches in view of vibroacoustic active control design.* PhD thesis, École Centrale Paris, 2009.

[43] E. Balmes, "A general component model synthesis method for damped structures," *AIAA Dynamics Specialist Conference, AIAA Paper 94-1692-CP*, pp. 182–190, 1994.

[44] E. Balmes, "Parametric families of reduced finite element models. theory and applications.," *International Modal Analysis Conference*, pp. 307–314, 1995.

[45] E. Balmes, "Optimal ritz vectors for component mode synthesis using the singular value decomposition," *AIAA Journal*, vol. 34, no. 6, pp. 1256–1260, 1996.

[46] E. Balmes, "Super-element representations of a model with frequency dependent properties," *International Seminar on Modal Analysis, Leuven, September*, vol. 3, pp. 1767–1778, 1996.

[47] E. Balmes, "Modes and regular shapes. how to extend component mode synthesis theory.," *Proceedings of the XI DINAME - Ouro Preto - MG - Brazil*, March 2005.

[48] H. Ben Dhia and E. Balmes, "Mesure de compatibilité et application aux problèmes de sous-structuration," *Colloque National en Calcul des Structures, Giens*, 2003.

[49] R. MacNeal, "A hybrid method of component mode synthesis," *Computers and structures*, vol. 1, no. 4, pp. 581–601, 1971.

[50] S. Rubin, "Improved component-mode representation for structural dynamic analysis," *AIAA Journal*, vol. 13, no. 8, pp. 995–1006, 1975.

[51] Y. Saad, *Iterative Methods for Sparse Linear Systems*. Self-published, second ed., January 2000.

[52] C. Ashcraft, *Ordering Sparse Matrices and Transforming Front Trees*. Boeing Shared Services Group, 1999.

[53] G. Karypis and V. Kumar, *METIS. A software package for partitioning unstructured graphs, partitioning meshes, and computing fill-reducing orderings of sparse matrices. Version 4.0*. Self published, September 1998.

[54] C. Ashcraft and R. Grimes, "Spooles: An object-oriented sparse matrix library," in *Proceedings of the 1999 SIAM Conference on Parallel Processing for Scientific Computing*, 1999.

[55] W. Gao, X. Li, C. Yang, and Z. Bai, "An implementation and evaluation of the amls method for sparse eigenvalue problems," *ACM Transactions on Mathematical Software*, vol. V, pp. 1–27, September 2007.

[56] B.-S. Liao, Z. Bai, and W. Gao, "The important modes subsystems: a moment matching approach," *Int. J. Numer. Meth. Engng.*, October 2006.

[57] D. C. Kammer and M. J. Triller, "Selection of component modes for craig-bampton substructure representations," *Journal of Vibration and Acoustics*, vol. 118, no. 2, pp. 264–270, 1996.

[58] D. Kammer and M. Triller, "Ranking the dynamic importance of fixed interface modes using a generalization of effective mass," *Modal Analysis: The International Journal of Analytical and Experimental Modal Analysis*, pp. 77–98, April 1994.

[59] I. Babuska, "The finite element method with lagrange multiplier," *Numer.Math.*, vol. 20, pp. 179–192, 1973.

[60] C. Bernadi, Y. Maday, and A. Patera, "Domain decomposition by the mortar element method," *Laboratoire d'analyse numérique, Paris VI Univ., Report 92013*, 1992.

[61] H.-G. Kim, "Interface element method (iem) for a partitioned system with non-matching interfaces," *Comput. Methods Appl. Mech. Engrg.*, vol. 1, no. 191, pp. 3165–3194, 2002.

[62] E. Balmes, "Use of generalized interface degrees of freedom in component mode synthesis," in *International Modal Analysis Conference*, pp. 204–210, 1996.

[63] E. Balmes, "Orthogonal maximum sequence sensor placements algorithms for modal tests, expansion and visibility.," *IMAC*, January 2005.

[64] W. Sextro, *Dynamic contact problems with friction. Models, Methods, Experiments and Applications*, vol. 3 of *Lecture Notes in Applied Dynamics*. Springer - Verlag Berlin Heidelberg New York, 2002.

[65] A. Sedaghat, J. Cooper, A. Leung, and J. Wright, "Estimation of the hopf bifurcation point for aeroelastic systems," *Journal of Sound and Vibration*, vol. 248, no. 1, pp. 31–42, 2001.

[66] J.-J. Sinou, O. Dereure, G.-B. Mazet, F. Thouverez, and J. Jezequel, "Friction-induced vibration for an aircraft brake system–part 1: Experimental approach and stability analysis," *International Journal of Mechanical Sciences*, vol. 48, no. 5, pp. 536–554, 2006.

[67] J.-J. Sinou, F. Thouverez, J. Jezequel, O. Dereure, and G.-B. Mazet, "Friction induced vibration for an aircraft brake system–part 2: Non-linear dynamics," *International Journal of Mechanical Sciences*, vol. 48, no. 5, pp. 555–567, 2006.

[68] S. Basseville and A. Leger, "Stability of equilibrium states in a simple system with unilateral contact and coulomb friction," *Archive of Applied Mechanics*, vol. 76, pp. 403–428, 2006.

[69] Q. S. Nguyen, "Instability and friction," *Comptes Rendus Mecanique*, vol. 331, no. 1, pp. 99–112, 2003.

[70] H. Hertz, "Über die berührung fester elastischer körper," *Journal für die reine und angewandte Mathematik (Crelles Journal)*, vol. 92, pp. 156–171, 1882.

[71] G. Amontons, "On the resistance originating in machines," *The French Royal Academy of Sciences*, pp. 206–220, 1699.

[72] C. Coulomb, *Theorie des machines simples en ayant égard au frottement des parties et à la roideur des cordes*. Mémoires de l'Académie des Sciences, 1821.

[73] R. Stribeck, "Die wesentlichen eigenschaften der gleit-und rollenlager," *Zeitschrift des Vereins Deutscher Ingenieure*, 1902.

[74] R. Mindlin, "Compliance of elastic bodies in contact," *J. of Applied Mechanics*, vol. 16, pp. 259–268, 1949.

[75] K. L. Johnson, "Surface interaction between elastically loaded bodies under tangential forces," *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, vol. 230, no. 1183, pp. pp. 531–548, 1955.

[76] J. F. Archard, "Elastic deformation and the laws of friction," *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, vol. 243, no. 1233, pp. pp. 190–205, 1957.

[77] A. Signorini, "Questioni di elasticita nonlinearizzata e semilinearizzata," *Rend. Mat.*, vol. 19, pp. 1–45, 1959.

[78] J. A. Greenwood and J. B. P. Williamson, "Contact of nominally flat surfaces," *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, vol. 295, no. 1442, pp. pp. 300–319, 1966.

[79] P. Painlevé, "Sur les lois du frottement de glissement," *CR Acad. Sci. Paris*, 1895.

[80] P. Ballard and S. Basseville, "Existence and uniqueness for dynamical unilateral contact with coulomb friction: a model problem," *M2AN*, vol. 39, no. 1, pp. 59–77, 2005.

[81] G. D. Saxcé and Z. Q. Feng, "The bipotential method: A constructive approach to design the complete contact law with friction and improved numerical algorithms," *Mathematical and Computer Modelling*, vol. 28, no. 4-8, pp. 225–245, 1998. Recent Advances in Contact Mechanics.

[82] J. Haslinger, "Least square method for solving contact problems with friction obeying the coulomb law," *Applications of Mathematics*, vol. 29, pp. 212–224, 1984.

[83] V. Acary, F. Cadoux, C. Lemaréchal, and J. Malick, "A formulation of the linear discrete coulomb friction problem via convex optimization," *Zeitschrift fur Angewandte Mathematik und Mechanik, to appear*, 2010.

[84] P. Alart and A. Curnier, "A mixed formulation for frictional contact problems prone to newton like solution methods," *Computer Methods in Applied Mechanics and Engineering*, vol. 92, no. 3, pp. 353–375, 1991.

[85] F. Jourdan, P. Alart, and M. Jean, "A gauss–siedel like algorithm to solve frictional contact problems," *Comput. Methods Appl. Mech. Engrg.*, vol. 155, pp. 31–47, 1998.

[86] H. Ben Dhia and M. Zarroug, "Hybrid frictional contact particles-in elements," *Revue Européenne des Eléments Finis*, vol. 11, no. 2-3-4, pp. 417–430, 2002.

[87] M. Jean, "The non-smooth contact dynamics method," *Computer methods in Applied Mechanics and Engineering*, no. 177, pp. 235–257, 1999.

[88] Z.-Q. Feng, M. Hjiaj, G. de Saxcé, and Z. Mróz, "Influence of anisotropy on contacting surfaces during loading/unloading cycles," *International Journal of Non-Linear Mechanics*, vol. 42, pp. 936–948, 2006.

[89] T. Baumberger and C. Caroli, "Solid friction from stick-slip to pinning and aging," *Advances in Physics*, vol. 55, pp. 279–348, May 2006.

[90] ABAQUS/Standard, *Theory Manual*, vol. 1. Hibbit, Karlsson, Sorhensen, Inc., 2008.

[91] M. Kham, "Elements de contact dérivés d'une formulation hybride continue," *Code_Aster, Version 8.5, R5.03.52-B*, pp. 1–50, 2007.

[92] D. Vola, M. Raous, and J. A. C. Martins, "Friction and instability of steady sliding: squeal of a rubber/glass contact," *Int. J. Numer. Meth. Engng.*, vol. 46, pp. 1699–1720, 1999.

[93] E. B. Pires and L. Trabucho, "The steady sliding problem with nonlocal friction," *International Journal of Engineering Science*, vol. 28, no. 7, pp. 631–641, 1990.

[94] L.-E. Andersson and A. Klarbring, "A review of the theory of static and quasi-static frictional contact problems in elasticity," *Phil Trans R Soc A*, vol. 359, pp. 2519–2539, 2001.

[95] A. M. Liapounoff, *Problème général de la stabilité du mouvement.* 2, 9, Annales de la faculté des sciences de Toulouse, 1907.

[96] J. J. Thomsen, *Vibration and stability. Advanced theory, analysis and tools*, vol. 3 of *Springer: Complexity.* Springer - Verlag Berlin Heidelberg New York, 2003.

[97] A. Lazarus and O. Thomas, "A harmonic-based method for computing the stability of periodic solutions of dynamical systems," *C. R. Mecanique*, vol. 338, pp. 510–517, 2010.

[98] W. Iwan, "A generalization of the concept of equivalent linearization," *International Journal of Non-Linear Mechanics*, vol. 8, no. 3, pp. 279–287, 1973.

[99] E. de Langre, "Frequency lock-in is caused by coupled-mode flutter," *Journal of Fluids and Structures*, vol. 22, no. 6-7, pp. 783–791, 2006.

[100] F. Cantone and F. Massi, "Effect of damping on the propensity of squeal instability: A numerical ivestigation," in *Proceedings of the International Conference on Advanced Acoustics and Vibration Engineering (ISMA)*, pp. 3603–3616, 2010.

[101] J. Carr, *Application of the center manifold theory*. Springer, New York, 1981.

[102] E. Pesheck, C. Pierre, and S. Shaw, "A new galerkin-based approach for accurate non-linear normal modes through invariant manifolds," *Journal of Sound and Vibration*, vol. 249, no. 5, pp. 971–993, 2002.

[103] J.-J. Sinou, F. Thouverez, and J. Jezequel, "Center manifold and multivariable approximants applied to non-linear stability analysis," *International Journal of Non-Linear Mechanics*, vol. 38, no. 9, pp. 1421–1442, 2003.

[104] A. Nayfeh and S. Nayfeh, "Nonlinear normal modes of a continuous system with quadratic nonlinearities," *Journal of Vibration and Acoustics*, vol. 117, no. 2, pp. 199–205, 1995.

[105] L. Hsu, "Analysis of critical and post-critical behaviour of non-linear dynamical systems by the normal form method, part i: Normalization formulae," *Journal of Sound and Vibration*, vol. 89, no. 2, pp. 169–181, 1983.

[106] L. Hsu, "Analysis of critical and post-critical behaviour of non-linear dynamical systems by the normal form method, part ii: Divergence and flutter," *Journal of Sound and Vibration*, vol. 89, no. 2, pp. 183–194, 1983.

[107] L. Jezequel and C. Lamarque, "Analysis of non-linear dynamical systems by the normal form theory," *Journal of Sound and Vibration*, vol. 149, no. 3, pp. 429 – 459, 1991.

[108] E. Petrov and D. Ewins, "Analytical formulation of friction interface elements for analysis of nonlinear multi-harmonic vibrations of bladed disks," *Journal of Turbomachinery*, vol. 125, no. 2, pp. 364–371, 2003.

[109] E. Petrov and D. Ewins, "Method for analysis of nonlinear multiharmonic vibrations of mistuned bladed disks with scatter of contact interface characteristics," *Journal of Turbomachinery*, vol. 127, no. 1, pp. 128–136, 2005.

[110] E. Petrov and D. Ewins, "Effects of damping and varying contact area at blade-disk joints in forced response analysis of bladed disk assemblies," *Journal of Turbomachinery*, vol. 128, no. 2, pp. 403–410, 2006.

[111] G. Von Groll and D. Ewins, "The harmonic balance method with arc-length continuation in rotor/stator contact problems," *Journal of Sound and Vibration*, vol. 241, no. 2, pp. 223–233, 2001.

[112] M. Guskov, J.-J. Sinou, and F. Thouverez, "Multi-dimensional harmonic balance applied to rotor dynamics," *Mechanics Research Communications*, vol. 35, no. 8, pp. 537–545, 2008.

[113] C. Pierre, A. Ferri, and E. Dowell, "Multi-harmonic analysis of dry friction damped systems using an incremental harmonic balance method," *Journal of Applied Mechanics*, vol. 52, no. 4, pp. 958–964, 1985.

[114] N. Coudeyras, J.-J. Sinou, and S. Nacivet, "A new treatment for predicting the self-excited vibrations of nonlinear systems with frictional interfaces: The constrained harmonic balance method, with application to disc brake squeal," *Journal of Sound and Vibration*, vol. 319, no. 3-5, pp. 1175–1199, 2009.

[115] N. Coudeyras, S. Nacivet, and J.-J. Sinou, "Periodic and quasi-periodic solutions for multi-instabilities involved in brake squeal," *Journal of Sound and Vibration*, vol. 328, no. 4-5, pp. 520–540, 2009.

[116] V. Jaumouillé, J.-J. Sinou, and B. Petitjean, "An adaptive harmonic balance method for predicting the nonlinear dynamic responses of mechanical systems–application to bolted structures," *Journal of Sound and Vibration*, vol. 329, no. 19, pp. 4048–4067, 2010.

[117] J.-J. Sinou, "Transient non-linear dynamic analysis of automotive disc brake squeal - on the need to consider both stability and non-linear analysis," *Mechanics Research Communications*, vol. 37, no. 1, pp. 96–105, 2010.

[118] R. Rosenberg, "On nonlinear vibrations of systems with many degrees of freedom," *Advances in Applied Mechanics*, vol. 9, pp. 155–242, 1966.

[119] S. W. Shaw and C. Pierre, "Normal modes for non-linear vibratory systems," *Journal of Sound and Vibration*, no. 1, pp. 85–124, 1993.

[120] C. Touzé, O. Thomas, and A. Chaigne, "Hardening/softening behaviour in non-linear oscillations of structural systems using non-linear normal modes," *Journal of Sound and Vibration*, vol. 273, no. 1-2, pp. 77–101, 2004.

[121] R. Seydel, "Nonlinear computation," *Journal of the Franklin Institute*, vol. 334, no. 5-6, pp. 1015–1047, 1997. Visions of Nonlinear Mechanics in the 21st Century.

[122] J. Ferreira and A. Serpa, "Application of the arc-length method in nonlinear frequency response," *Journal of Sound and Vibration*, vol. 284, no. 1-2, pp. 133–149, 2005.

[123] B. Cochelin and C. Vergez, "A high order purely frequency-based harmonic balance formulation for continuation of periodic solutions," *Journal of Sound and Vibration*, vol. 324, no. 1-2, pp. 243–262, 2009.

[124] J.-J. Sinou, F. Thouverez, and L. Jezequel, "Stability analysis and non-linear behaviour of structural systems using the complex non-linear modal analysis (cnlma)," *Computers & Structures*, vol. 84, no. 29-30, pp. 1891–1905, 2006.

[125] G. Kerchen, M. Peeters, J.-C. Golinval, and A. Vakakis, "Nonlinear normal modes, part i: A useful framework for the structural dynamicist," *Mechanical Systems and Signal Processing*, no. 23, pp. 170–194, 2009.

[126] M. Peeters, R. Viguié, G. Sérandour, G. Kershen, and J.-C. Golinval, "Nonlinear normal modes, part ii: Toward a practical computation using numerical continuation techniques," *Mechanical Systems and Signal Processing*, no. 23, pp. 195–216, 2009.

[127] E. Sarrouy and F. Thouverez, "Global search of non-linear systems periodic solutions: A rotordynamics application," *Mechanical Systems and Signal Processing*, vol. 24, no. 6, pp. 1799–1813, 2010.

[128] C. S. Hsu, "A theory of cell-to-cell mapping dynamical systems," *Journal of Applied Mechanics*, vol. 47, no. 4, pp. 931–939, 1980.

[129] C. S. Hsu and B. H. Tongue, "Cell-to-cell mapping, a method of global analysis for nonlinear systems," *Journal of Applied Mechanics*, vol. 55, no. 3, pp. 749–750, 1988.

[130] B. Tongue, "On obtaining global nonlinear system characteristics through interpolated cell mapping," *Physica D: Nonlinear Phenomena*, vol. 28, no. 3, pp. 401–408, 1987.

[131] K. Georg, "A new exclusion test," *Journal of Computational and Applied Mathematics*, vol. 152, no. 1-2, pp. 147–160, 2003. Proceedings of the International Conference on Recent Advances in Computational Mathematics.

[132] N. Newmark, "A method of computation for structural dynamics," *J. Engng Mech. Div., ASCE 85*, pp. 67–94, 1959.

[133] J. Houbolt, "A recurrence matrix solution for the dynamic response of elastic aircraft," *Journal of the Aeronautical science*, no. 17, pp. 540–550, 1950.

[134] O. Zienkiewicz, "A new look at the newmark, houbolt and other time stepping formulas. a weighted residual approach," *Earthquake Engineering & Structural Dynamics*, vol. 5, no. 4, pp. 413–418, 1977.

[135] K. Bathe and E. Wilson, "Stability and accuracy analysis of direct integration methods," *Earthquake Engineering and Structural Dynamics*, vol. 1, pp. 283–291, 1973.

[136] M. Hilber, T. Hughes, and R. Taylor, "Improved numerical dissipation for time integration algorithms in structural dynamics," *Earthquack Engineering & Structural Dynamics*, vol. 5, no. 3, pp. 283–292, 1977.

[137] S. Krenk, "State-space time integration with energy control and fourth-order accuracy for linear dynamic systems," *Int. J. Numer. Meth. Engng.*, vol. 65, pp. 595–619, 2006.

[138] H. Kehnous, P. Laborde, and Y. Renard, "Mass redistribution method for finite element contact problems in elastodynamics," *European Journal of Mechanics A*, no. 27, pp. 918–932, 2008.

[139] M. Géradin and D. Rixen, *Mechanical Vibrations. Theory and Application to Structural Dynamics.* John Wiley & Wiley and Sons, 1994, also in French, Masson, Paris, 1993.

[140] B. Tie and D. Aubry, "Reliable and efficient numerical modeling of elastic wave propagation in engineering structures by discontinuous galerkin fem," in *Computational Mechanics, WCCM VI in conjunction with APCOM'04*, September 2004.

[141] A. Grédé, *Modélisation des chocs d'origine pyrotechnique dans les structures d'Ariane5 : développement de modèles de propagation et d'outils de modélisation.* PhD thesis, École Centrale Paris, 2009.

[142] H. Adouani, B. Tie, C. Berdin, and D. Aubry, "Adaptative numerical modeling of dynamic crack propagation," *J. Phys. IV. France*, vol. 134, pp. 501–506, August 2006.

[143] M. Mancuso and F. Ubertini, "An efficient time discontinuous galerkin procedure for nonlinear structural dynamics," *Comput. Methods Appl. Mech. Engrg.*, vol. 195, pp. 6391–6406, 2006.

[144] M. Hulbert and T. Hughes, "Space-time finite element methods for second-order hyperbolic equations," *Computer methods in Applied Mechanics and Engineering*, no. 84, pp. 327–348, 1990.

[145] C. Chien, C. Yang, and J. Tang, "Three-dimensional transient elastodynamic analysis by a space and time-discontinuous galerkin finite element method," *Finite Elements in Analysis and Design*, vol. 39, pp. 561–580, 2003.

[146] G. Hulbert and T. Hughes, "Space-time finite element methods for second-order hyperbolic equations," *Computer Methods in Applied Mechanics and Engineering*, vol. 84, no. 3, pp. 327–348, 1990.

[147] A. Bonelli, O. Bursi, and M. Mancuso, "Explicit predictor-multicorrector time discontinuous galerkin methods for non-linear dynamics," *Journal of Sound and Vibration*, vol. 4, pp. 695–724, 2002.

[148] M. Raous, S. Barbarin, and D. Vola, "Numerical characterization and computation of dynamic instabilities for frictional contact problems in friction and instabilities," *CSIM, Courses ans Lectures*, pp. 234–291, 2002.

[149] T. Hughes, "Stability, convergence and growth and decay of energy of the average acceleration method in nonlinear structural dynamics," *Computers & Structures*, vol. 6, pp. 313–324, 1976.

[150] M. Géradin and A. Cardona, "Time integration of the equations of motion in mechanism analysis," *Computers & Structures*, vol. 33, no. 3, pp. 801–820, 1989.

[151] A. Mugan, "Frequency domain analysis of time integration algorithms," *Comput. Methods Appl. Mech. Engrg*, vol. 190, pp. 5777–5793, 2001.

[152] J. Chung and G. M. Hulbert, "A time integration algorithm for structural dynamics with improved numerical dissipation: The generalized-alpha method," *Journal of Applied Mechanics*, vol. 60, no. 2, pp. 371–375, 1993.

[153] O. Brüls and M. Arnold, "The generalized-alpha scheme as a linear multistep integrator: Toward a general mechatronic simulator," *Journal of Computational and Nonlinear Dynamics*, vol. 3, no. 4, p. 041007, 2008.

[154] C.-A. Felippa and K.-C. Park, "Direct time integration methods in non linear structural dynamics," *Comput. Methods Appl. Mech. Engrg*, vol. 17/18, pp. 277–313, 1979.

[155] P. Badel, J. Laverne, and N. Tardieu, "Algorithme non linéaire quasi-statique," *Code_Aster, Version 7.4, R5.03.01-D*, pp. 1–28, 2005.

[156] A. Nashif, D. Jones, and J. Henderson, *Vibration Damping*. John Wiley and Sons, 1985.

[157] E. Balmes, *Viscoelastic vibration toolbox, User Manual*. SDTools, 2004-2007.

[158] B. D. Coleman and W. Noll, "Foundations of Linear Viscoelasticity," *Reviews of Modern Physics*, vol. 33, pp. 239–249, 2 1961.

[159] J. J. Espindola, J. M. Silva Neto, and E. M. O. Lopes, "A new approach to viscoelastic material properties identification based on the fractional derivative model," *Proceedings of First IFAC Workshop on Fractional Differentiation and its Application, Bordeaux*, pp. 271–276, July 2004.

[160] J. Ferry, *Viscoelastic Properties of Polymers*. Wiley, 2nd ed., 1970.

[161] S. Park, "Analytical modeling of viscoelastic dampers for structural and vibration control," *International Journal of Solids and Structures*, vol. 38, pp. 8065–8092, 2001.

[162] L. Rogers, C. Johson, and D. Keinholz, "The modal strain energy finite element method and its application to damped laminated beams," *Shock and Vibration Bulletin*, vol. 51, 1981.

[163] T. Hasselman, "Modal coupling in lightly damped structures," *AIAA Journal*, vol. 14, no. 11, pp. 1627–1628, 1976.

[164] T. Caughey, "Classical normal modes in damped linear dynamic systems," *ASME J. of Applied Mechanics*, pp. 269–271, 1960.

[165] E. Balmes, "New results on the identification of normal modes from experimental complex modes," *Mechanical Systems and Signal Processing*, vol. 11, no. 2, pp. 229–243, 1997.

[166] J.-P. Bianchi, E. Balmes, G. Vermot des Roches, and A. Bobillot, "Using modal damping for full model transient analysis," in *Proceedings of the International Conference on Advanced Acoustics and Vibration Engineering (ISMA)*, 2010.

[167] G. Strang, *Introduction to linear algebra*. Wellesley Cambridge Press, 4^{th} edition, 2009.

[168] G. Fritz, J.-J. Sinou, M. Duffal, and L. Jezequel, "Investigation of the relationship between damping and mode coupling patterns in case of brake squeal," *Journal of Sound and Vibration*, vol. 307, pp. 591–609, 2008.

[169] B. Hervé, J.-J. Sinou, H. Mahé, and L. Jézéquel, "Extension of the destabilization paradox to limit cycle amplitudes for a nonlinear self-excited system subject to gyroscopic and circulatory actions," *Journal of Sound and Vibration*, vol. 323, no. 3-5, pp. 944–973, 2009.

[170] J.-J. Sinou and L. Jézéquel, "The influence of damping on the limit cycles for a self-exciting mechanism," *Journal of Sound and Vibration*, vol. 304, no. 3-5, pp. 875–893, 2007.

[171] E. Anderson, Z. Bai, C. Bischof, L. S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenny, and D. Sorensen, *Lapack Users' Guide, third edition*. SIAM, www.netlib.org/lapack, 2000.

[172] K.-J. Bathe and E. L. Wilson, "Solution methods for eigenvalue problems in structural mechanics," *International Journal for Numerical Methods in Engineering*, vol. 6, pp. 213–226, 1973.

[173] W. E. Arnoldi, "The principle of minimized iterations in the solution of the matrix eigenvalue problem," *Quart. Appl. Meth.*, pp. 17–19, 1951.

[174] D. C. Sorensen, "Implicit application of polynomial filters in a k-step arnoldi method," *SIAM Journal on Matrix Analysis and Applications*, vol. 13, no. 1, pp. 357–385, 1992.

[175] P. R. Dahl, "Solid friction damping of mechanical vibrations," *AIAA Journal*, vol. 14, pp. 1675–1682, 1976.

[176] G. Chevallier, *Etude des vibrations de broutement provoquées par le frottement sec. Application aux systèmes d'embrayage*. PhD thesis, Université Paris VI Pierre et Marie Curie, 2005.

[177] B. Armstrong-Hélouvry, P. Dupont, and C. Canudas De Wit, "A survey of models, analysis tools and compensation methods for the control of machines with friction," *Automatica*, vol. 30, no. 7, pp. 1083–1138, 1994.

[178] B. Borsotto, *Modélisation, Identification et commande d'un organe de friction. Application au contrôle d'un système d'embrayage et au filtrage d'acyclismes par glissement piloté.* PhD thesis, Supélec, 2008.

[179] D. Karnopp, "Computer simulation of stick-slip friction in mechanical dynamic systems," *Journal of Dynamic Systems, Measurement, and Control*, vol. 107, no. 1, pp. 100–103, 1985.

[180] F. Al-Bender, V. Lampaert, and J. Swevers, "The generalized maxwell-slip model: a novel model for friction simulation and compensation," *Transactions on Automatic Control, IEEE*, vol. 50, pp. 1883–1887, 2005.

[181] J. Rice and A. Ruina, "Stability of steady frictional slipping," *J. Appl. Mech.*, vol. 50, pp. 343–349, 1983.

[182] C. Canuda de Wit, H. Olsson, K. Astrôm, and P. Lischinsky, "A new model for control of systems with friction," *IEEE transactions on automatic control*, vol. 40, no. 3, pp. 419–425, 1995.

[183] V. Lampaert, F. Al-Bender, and J. Swevers, "Experimental characterization of dry friction at low velocities on a developed tribometer setup for macroscopic measurements," *Tribology Letters*, vol. 16, pp. 95–105, 2004.

[184] J. Swevers, F. Al-Bender, C. Ganseman, and T. Projogo, "An integrated friction model structure with improved presliding behavior for accurate friction compensation," *Transactions on Automatic Control, IEEE*, vol. 45, pp. 675–686, 2000.

[185] P. Bliman and M. Sorine, "Easy-to-use realistic dry friction models for automatic control," in *Proceedings of 3rd European Control Conference*, pp. 3788–3794, 1995.

[186] D. Haessig and B. Friedland, "On the modeling and simulation of friction," *Journal of Dynamic Systems, Measurement and Control, Transactions ASME*, 1991.

[187] P. Dupont, B. Armstrong, and V. Hayward, "Elasto-plastic friction model: contact compliance and stiction," *Proc. Amer. Control Conf., Chicago, II*, 2000.

List of Figures

List of Tables