Numerical methods for parametric model reduction in the simulation of disk brake squeal [☆]

Nils Gräbner[b], Volker Mehrmann[a], Sarosh Quraishi[a], Christian Schröder[a], Utz von Wagner[b]

[a]*Department of Mathematics, Chair of Numerical Mathematics, Technische Universität Berlin, Sekr. MA 4-5, Strasse des 17. Juni 136, 10623 Berlin, Germany*
[b]*Department of Applied Mechanics, Chair of Mechatronics and Machine Dynamics, Technische Universität Berlin, Sekr. MS1, Einsteinufer 5, 10587 Berlin, Germany*

Abstract

We present numerical methods for model reduction in the numerical simulation of disk brake squeal. Automotive disk brake squeal is a high frequency noise phenomenon based on self excited vibrations. Our method is based on a variation of the proper orthogonal decomposition method and involves the solution of a large scale, parametric eigenvalue problem. Several important challenges arise, some of which can be traced back to the finite element modeling stage. Compared to the current industrial standard our new approach is more accurate in vibration prediction and achieves a better reduction in model size. This comes at the price of an increased computational cost, but it still gives useful results when the traditional method fails to do so. We illustrate the results with several numerical experiments, some from real industrial models, some from simpler academic models. These results indicate where improvements of the current black box industrial codes are advisable.

Keywords: brake squeal, quadratic eigenvalue problem, complex eigenvalue analysis, model reduction, damped systems, modeling errors, proper orthogonal decomposition
2010 MSC: 65F15, 65N30, 65P40

1. Introduction

Brake squeal of disk brakes is a noise pollution problem that affects people to a large extent. It can be interpreted as noise generated due to self-excited vibration caused by a flutter-type instability originating from friction forces at the pad-rotor interface [1]. Due to friction induced forces the whole brake system (see Fig. 1.1 for a general view) oscillates in a vibration mode, ultimately reaching a limit-cycle. Several reasons have been suggested for the onset of the instabilities. The analysis is usually based on lab experiments, on numerical simulations based on finite element models, or on idealized minimal models mimicking the physics of a real brake. A broad overview can be found in [25], whereas numerical aspects are reviewed in [33]. Other aspects of disk brake squeal have been considered in [10, 11, 20, 31]. Despite extensive research, however, the problem is far from being understood and, so far, satisfactory remedies have not been found that can be implemented in a systematic way.

The macroscopic equations of motion arising from finite element modeling are usually considered in the form

$$M_\Omega \ddot{u} + D_\Omega \dot{u} + K_\Omega u = f, \tag{1.1}$$

where M_Ω, D_Ω, and $K_\Omega \in \mathbb{R}^{n,n}$ are parameter–dependent coefficient matrices collecting acceleration, velocity, and displacement proportional terms, respectively, and f is an external force. In

particular, M_Ω is the mass matrix, D_Ω collects damping and gyroscopic effects, and K_Ω collects stiffness and circulatory effects. The vector valued function $u : \mathbb{R} \to \mathbb{R}^n$ contains the coordinates (in the FE basis) of the displacements, while the time derivatives $\dot{u}$, $\ddot{u}$ contain the corresponding coordinates of the velocity and acceleration. The coefficient matrices are usually real, large, and sparsely populated with nonzero entries. M_Ω is symmetric and positive semidefinite. For FE models of rotating machinery (such as a disk brake), D_Ω and K_Ω are typically nonsymmetric due gyroscopic and circulatory forces. For self-excited vibrations, which are of interest here, we have $f \equiv 0$. All three matrices may depend on one or more parameters, here denoted by the parameter vector Ω. These typically include operating conditions (temperature, pad pressure, etc), material properties (friction coefficient, brake geometry and mass distribution, effects of wear and damping etc) as well as the rotational speed of the brake disk. In all our numerical examples we vary only a single parameter, the rotational speed of the brake disk $\Omega \in [\Omega_{min}, \Omega_{max}] \subset \mathbb{R}$.

Assuming the linear model (1.1), the dynamical behavior of the system and, in particular, the excitation of vibrational modes can be investigated based on the eigenvalues of the quadratic eigenvalue problem (QEP)

$$(\lambda^2 M_\Omega + \lambda D_\Omega + K_\Omega)x = 0. \tag{1.2}$$

This analysis, which is popularly known as complex eigenvalue analysis (CEA), is also valid locally, when a nonlinear model is linearized around a stationary equilibrium solution [36], however, it is usually not valid when the system is nonlinear and the linearization is around a non-stationary limit cycle solution [3]. According to classical linear stability analysis (see, e.g., [14, 16, 41]) the equilibrium is stable when the real parts of all eigenvalues of the linearized system are negative. When at least one real part is positive or there is a multiple purely imaginary eigenvalue then the equilibrium becomes unstable – independently of nonlinear effects.

Thus, in order to analyze the stability of the FE brake model (Fig. 2.1), the eigenvalues of the nonsymmetric quadratic eigenvalue problem (1.2) in the right half plane, in the following called unstable eigenvalues, have to be determined. There exist a multitude of numerical methods for the numerical solution of quadratic eigenvalues problems, see e.g. [32, 38]. However, solving (1.2) still presents a major computational challenge, in particular, since this has to be done for many values of the parameter Ω.
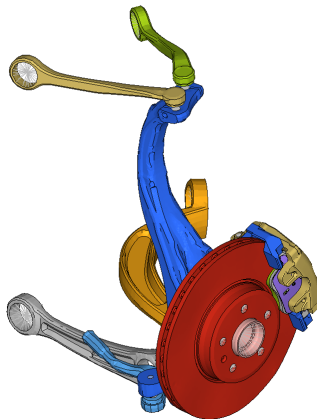

Figure 1.1: General view of the brake model with adjacent components

In this paper we discuss a parametric model reduction technique that produces from the large scale system (1.2) a small scale system in which the unstable behavior of the system is present. Hence, the small system may be solved cheaply for many parameter values and can thus be used for optimizing the original system with respect to parameter variations very efficiently. More precisely, we construct small scale parametric matrices $\tilde{M}_\Omega, \tilde{D}_\Omega, \tilde{K}_\Omega \in \mathbb{R}^{d,d}$ with $d \ll n$ such that for all parameters $\Omega \in [\Omega_{min}, \Omega_{max}]$ the unstable eigenvalues of $(\tilde{M}_\Omega, \tilde{D}_\Omega, \tilde{K}_\Omega)$ are close to the unstable eigenvalues of $(M_\Omega, D_\Omega, K_\Omega)$. Moreover, still for all parameters $\Omega \in [\Omega_{min}, \Omega_{max}]$, if (λ, x) is an

unstable eigenvalue/eigenvector of $(M_\Omega, D_\Omega, K_\Omega)$ then there is an eigenvalue/eigenvector $(\tilde{\lambda}, \tilde{x})$ of $(\tilde{M}_\Omega, \tilde{D}_\Omega, \tilde{K}_\Omega)$ such that $\tilde{\lambda} \approx \lambda$ and $\phi(\tilde{x}) \approx x$. Here ϕ denotes a prolongation mapping $\mathbb{R}^d$ to $\mathbb{R}^n$.

We will construct the small scale system by projection to an appropriate subspace. This general approach is particularly useful for the brake squeal problem, since brake squeal is typically a mono-frequent phenomenon, where only a single vibration mode dominates. In these cases an ideal projection space is obtained by applying the method of proper orthogonal decomposition (POD) [24, 42] to a few eigenvectors corresponding to the squeal frequency (even when varying the parameters). In the discussed examples the dimension of the parametric FE model can be reduced from around one million to about fifty, while retaining good accuracy of the solutions (λ, x) for the whole parameter set.

While applying our new technique to an industrial brake model, we noticed severe sensitivity issues of the eigenvalues and eigenvectors under small perturbations. Actually we observed that the QEP (1.2) was often close to a singular problem. We analyzed the FE models and detected potential reasons for the high sensitivity. In order to detect the presence of these potential difficulties due to improper modeling and also to deflate singular parts in the FE model, we derive some techniques to analyze FE models and to avoid the singularities.

The paper is organized as follows. In Section 2 we recall the standard modeling procedures for disk brakes. The resulting quadratic eigenvalue problem is discussed in Section 3. We review the traditional model reduction for parametric eigenvalue problems in Section 4. Then, in Section 5 we present the POD approach. Numerical results in using industrial disk brake models with a large number of degrees of freedom are presented in Section 6, where we also compare our method with the classical model reduction techniques. We present some concluding remarks in Section 8. Numerical procedures for the solution of large scale quadratic eigenvalue problems are discussed in Section Appendix A. Finally, in Appendix B we discuss the difficulties arising from almost singular eigenvalue problems and we show that this may arise from improper FE modeling.

2. Derivation of the disk brake model

Complicated dynamical systems such as disk brakes can be considered as a collection of structural elements considered to be continua in mechanical modeling. Fundamentally, at the micro (atomistic) scale this description is based on the Langevin equation (e.g., [35]). Unfortunately, full atomistic simulation of structures like a disk brake is not computationally feasible. So one resorts to approximate representations at a macro scale, e.g., by using multibody dynamics and finite element modeling. But also at the macro scale it is very difficult to construct universally valid phenomenological models of material damping, internal friction and friction forces, [12, 13].

Thus, although nonlinearities need to be studied and incorporated into the model for a *true* understanding of their nature [11], useful results can be obtained using a linearized finite element (FE) model, formulating the equations of motion assuming a very simplied description of the forcing term arising from a macroscopic friction law, e.g., the Coulomb law [34].

2.1. FE modeling

This section summarizes the FE modeling methods for brake squeal as they are implemented in software packages such as [21]. The goal of FE modeling is to develop a mathematical model which is able to describe the mechanical characteristics of the physical system sufficiently well to be able to derive conclusions about the behavior of the physical system. For the stability analysis of the disk brake model this is done via the computation of eigenvalues and eigenmodes. In linear stability analysis, if there exist eigenvalues with positive real part, then a self-excited friction-induced vibration may arise and for real brakes this oscillation emerges as a potentially audible squeal.
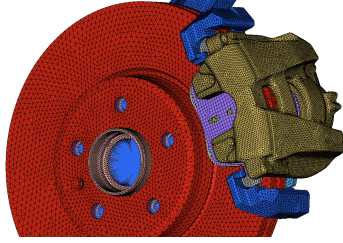
Figure 2.1: example of a FE-Model of a disk brake

At the beginning of the analysis, the initial state of the brake is that of an unloaded stationary brake. At this point all of the possible contact zones are defined but not in contact yet, and furthermore, the rotation of the disk is neglected as well. To consider all effects which result from applying brake pressure and thus from normal and friction forces it is essential to go through several steps of model refinement.

Neglecting external excitation the equation of motion has the form

$$M\ddot{u} + D_M\dot{u} + K_E u = 0,$$

where M, D_M and K_E are symmetric mass, damping (material damping), and elastic stiffness matrices, M and K_E are positive definite and D_M is positive semidefinite.

2.2. Linear static analysis to determine the contact state

In a first step we perform a linear static analysis. The system under investigation is the disk with the external load from the brake pad. Although the disk is considered stationary, we assign velocity field information to each FE node to map the friction forces at the contact correctly. The linear static analysis provides the location of contact as well as the normal and friction forces in the contact area (using the Coulomb model of friction [4, 34]).
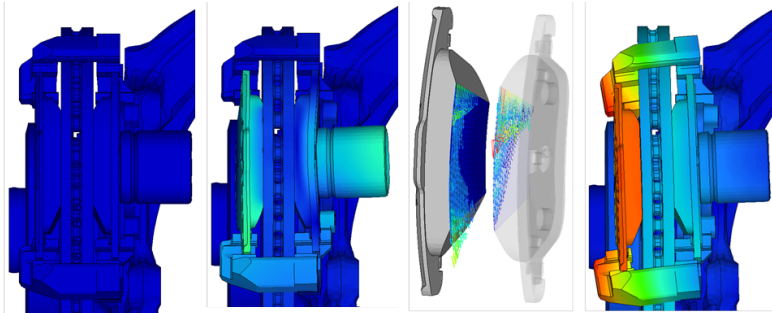

Figure 2.2: (From left to right) unloaded system, contact forces in normal direction due to the brake pressure, friction forces on the brake pad, deformation caused by friction forces

To refine this model further, the state of contact is frozen and the contact points are constrained in normal direction with multi-point constraints (MPCs). This extends the equations of motion by two additional terms $K_R u$ and $\frac{1}{\Omega}D_R\dot{u}$, where Ω is the rotational speed of the disk. Here K_R is nonsymmetric and describes circulatory effects and the term D_R is symmetric and describes the friction-induced damping. The extended equations of motion then have the form

$$M\ddot{u} + (D_M + \frac{1}{\Omega}D_R)\dot{u} + (K_E + K_R)u = f. \tag{2.1}$$

4

2.3. Linear static analysis with centrifugal loads

In the next step the modified model (2.1) is loaded as if the disk brake was rotating. But instead of moving the nodes, they are applied with the load resulting from centrifugal forces, see Fig. 2.3.
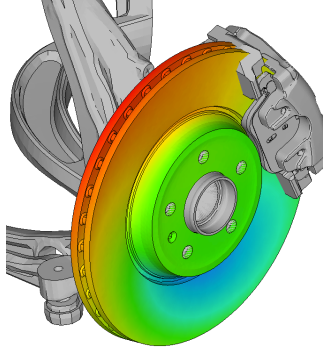

Figure 2.3: Updated model including MPCs and centrifugal loads resulting from a rotation of the brake disk

With this analysis it is possible to determine the internal stress conditions. This procedure leads to additional terms $\Omega^2 K_g u$ and $\Omega D_G \dot{u}$ in the model. The extended equations of motion (in homogeneous form) then are

$$M\ddot{u} + (D_M + \frac{1}{\Omega}D_R + \Omega D_G)\dot{u} + (K_E + K_R + \Omega^2 K_g)u = 0, \tag{2.2}$$

where D_G is a skew-symmetric matrix (arising from the modeling of the gyroscopic terms) and K_g is a symmetric matrix modeling the geometric stiffness. These coefficient matrices incorporate all effects of the linear static analysis. Note that the parameter dependence is expressed only in the disk speed Ω. In general, all the coefficient matrices may also depend on other parameters, but this dependence is not considered here.

3. Eigenvalue analysis

We review the classical analysis connecting instability to eigenvalues. Disc brake squeal is usually assumed to be caused by an instability of the equilibrium solution $u(t) \equiv 0$ of (2.2). Using the ansatz $u(t) = \text{Re}(e^{\lambda t}x)$ we obtain the parameter dependent quadratic eigenvalue problem (QEP)

$$P_\Omega(\lambda_\Omega)x_\Omega = 0, \quad \text{where} \quad P_\Omega(\lambda) = \lambda^2 M + \lambda D_\Omega + K_\Omega, \tag{3.1}$$

where D_Ω and K_Ω denote the collected coefficient matrices of (2.2). As discussed before, the parameter Ω here denotes the angular velocity of the disk (measured in radians per second).

To analyze when a disk brake is expected to squeal, we need to compute the (typically few) eigenvalues with positive real part of (3.1), because these are often used as a criterion for brake squeal [28]. In particular, the magnitude of the positive real part of an eigenvalue is correlated with the tendency of a brake to squeal.

The classic approach to solve a QEP is to first perform a first order formulation, also known as *linearization* [7]. It transforms the QEP to an equivalent linear eigenvalue problem of double dimension with the same eigenvalues. There are many different linearizations and they have been studied in detail with respect to their properties, see [19, 29]. The classical companion linearization of (3.1) has the form

$$\left(\lambda \begin{bmatrix} M & 0 \\ 0 & I \end{bmatrix} + \begin{bmatrix} D_\Omega & K_\Omega \\ -I & 0 \end{bmatrix} \right) \begin{bmatrix} \lambda x \\ x \end{bmatrix} = 0. \tag{3.2}$$

If the mass matrix M is positive definite (which is the case when no position or velocity constraints are modeled directly) and also the part $K_E + \Omega^2 K_g$ of the stiffness matrix is positive definite (which should be the case in a reasonable FE model) then we can scale and write (3.2) in the form

$$(\lambda \hat{M} + (\hat{G} + \hat{D}_1 + \hat{D}_2))y = 0,$$

with a symmetric positive definite matrix $\hat{M}$, a skew-symmetric matrix $\hat{G}$, a symmetric positive semidefinite matrix $\hat{D}_1$, and a symmetric indefinite matrix $\hat{D}_2$:

$$\hat{M} := \begin{bmatrix} M & 0 \\ 0 & K_E + \Omega^2 K_g \end{bmatrix}, \qquad \hat{G} := \begin{bmatrix} \Omega D_G & K_E + \Omega^2 K_g + \frac{1}{2} K_R \\ -(K_E + \Omega^2 K_g + \frac{1}{2} K_R)^T & 0 \end{bmatrix},$$

$$\hat{D}_1 := \begin{bmatrix} D_M + \frac{1}{\Omega} D_R & 0 \\ 0 & 0 \end{bmatrix}, \qquad \hat{D}_2 := \begin{bmatrix} 0 & \frac{1}{2} K_R \\ \frac{1}{2} K_R^T & 0 \end{bmatrix}, \quad y := \begin{bmatrix} \lambda x \\ x \end{bmatrix}.$$

In this form we immediately observe that (independently of Ω) all eigenvalues of $\lambda \hat{M} + \hat{G}$ are purely imaginary, they are the eigenvalues of the skew-symmetric matrix $-L^{-1} \hat{G} L^{T-1}$, where $\hat{M} = LL^T$ is the Cholesky decomposition [9]. Adding the semidefinite block matrix $\hat{D}_1$ moves the spectrum to the left in the complex half plane, so that $\hat{D}_2$, which is resulting from the circulatory effects, is solely responsible for moving eigenvalues to the right half plane.

This simple analysis shows that eigenvalue analysis allows to study the effects of (friction) induced oscillations — provided that M and $K_E + \Omega^2 K_g$ are positive definite. It should be noted however, that this analysis has to be modified if M or $K_E + \Omega^2 K_g$ are only semidefinite.

A discussion of numerical methods to solve the quadratic eigenvalue problem is presented in Appendix A.

4. Model reduction, traditional approach

The standard approach for reducing large-scale (n dimensional) parametric quadratic eigenvalue problems $P_\Omega(\lambda)x = 0$, is to project the problem into a much smaller (say, d dimensional with $d << n$) subspace $\mathcal{Q}$, represented as the range of a rectangular matrix $Q \in \mathbb{R}^{n,d}$. In the case of parametric eigenvalue problems, this subspace should be independent of the parameter Ω. The reduced QEP in the disk brake problem then takes the form

$$\begin{aligned} P_\Omega^r(\lambda)\, x^r = 0, \quad \text{with} \quad P_\Omega^r(\lambda) &= Q^T P_\Omega(\lambda) Q \\ &= Q^T (\lambda^2 M + \lambda D_\Omega + K_\Omega)\, Q. \end{aligned}$$

Introducing the $d \times d$ matrices

$$M^r = Q^T M Q, \qquad\qquad D_\Omega^r = Q^T D_\Omega Q, \qquad\qquad K_\Omega^r = Q^T K_\Omega Q,$$

we have the reduced QEP

$$P_\Omega^r(\lambda_\Omega)x_\Omega^r = (\lambda_\Omega^2 M^r + \lambda_\Omega D_\Omega^r + K_\Omega^r)x_\Omega^r = 0.$$

This reduced model can be completely solved for any given value of Ω by a full dense eigenvalue method such as, e.g., `quadeig` [15] or MATLAB's `polyeig` [30]. Approximations of an eigenvector x_Ω of the large scale problem are then given by Qx_Ω^r.

It remains to choose an appropriate projection space $\mathcal{Q}$. Let us consider a simple eigenvalue λ and associated eigenvector x of P_Ω in the open right half complex plane for some value of Ω. Since eigenvalues depend continuously on the coefficient matrices [9], it follows that P_Ω^r has an eigenvalue close to λ, whenever x is approximately contained in the subspace $\mathcal{Q}$. Thus, $\mathcal{Q}$ should contain good approximations to the eigenvectors x_Ω of the original matrix polynomial $P_\Omega(\lambda)$ associated with the right half plane eigenvalues for all values of the parameter Ω.

The traditional choice of Q is presented in the next subsection. Our new method differs from the traditional approach in the choice of Q.

6

4.1. Traditional choice of the projection space

The traditional approach (often called modal truncation), that is used in commercial FE packages such as [21], chooses $\mathcal{Q}$ as the space spanned by the eigenvectors $x_1, \ldots, x_d$ associated with the d smallest eigenvalues of the linear symmetric eigenvalue problem

$$K_{\text{ref}} x = \mu M x, \tag{4.1}$$

where $K_{\text{ref}} := K_E + \Omega_{\text{ref}}^2 K_g$ for some reference value $\Omega_{\text{ref}} \in [\Omega_{min}, \Omega_{max}]$, i.e., $Q = [x_1, \ldots, x_d]$.

The eigenvalue problem (4.1) corresponds to the simplified dynamical system

$$M \ddot{u} + (K_E + \Omega_{\text{ref}}^2 K_g) u = f, \tag{4.2}$$

that is obtained from (2.2) by ignoring the parameter dependent damping terms and the effect of friction forces, and setting Ω to Ω_{ref}. However, the MPCs which connect the contact points of the disk and the pad are incorporated in (4.2). If M and K_{ref} are positive definite, then the associated matrix polynomial $\lambda^2 M + K_{\text{ref}}$ has only purely imaginary eigenvalues, or alternatively the matrix pencil $-\mu M + K_{\text{ref}}$ has only real and positive eigenvalues. This follows by setting $\lambda = \imath \sqrt{\mu}$ [26]. This means that the dynamical system (4.2) is stable but not asymptotically stable. If M is only semidefinite, which is the case if rigid connections are included in the model, then there also exist infinite eigenvalues. If K_{ref} is singular then there are also zero eigenvalues. Note that it may also happen that $-\mu M + K_{\text{ref}}$ is close to singular for almost all μ which indicates bad FE modeling.

The traditional modal truncation approach can be applied successfully in many cases. Fig. 4.1, e.g., shows an unstable mode shape obtained by this approach that looks reasonable.
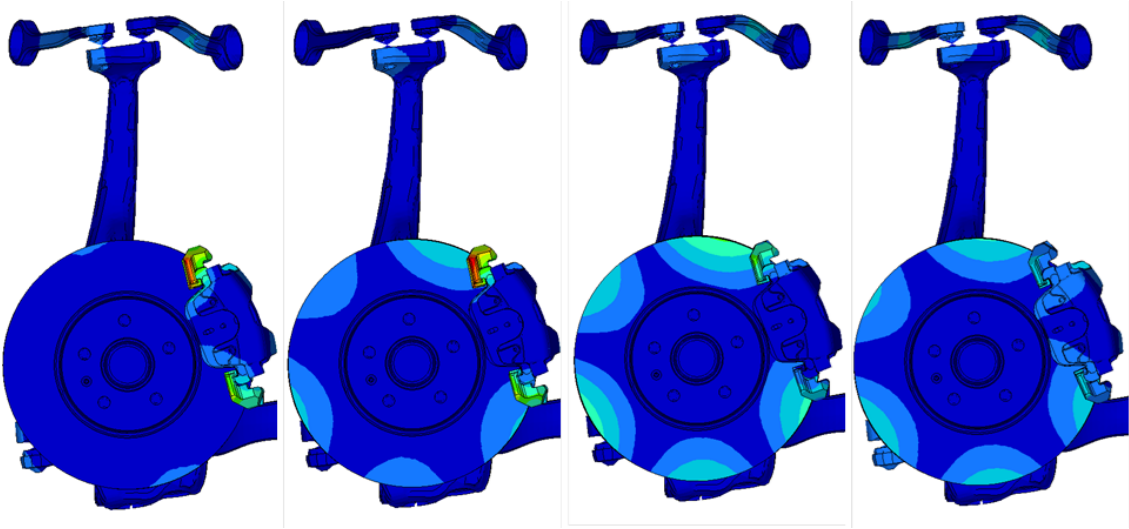

Figure 4.1: Complex eigenform associated with positive real part eigenvalue, $\text{Im}(x e^{\imath \phi})$ for a phase of $\phi = 0, 45, 90,$ and 135 degrees at $\Omega = 1873$ Hz.

However, for the modal truncation approach to be justified, the influence of the omitted parameter dependent terms must be small. But in some case, as we will demonstrate below, it may completely miss important effects.

5. A proper orthogonal decomposition approach

Proper orthogonal decomposition (POD) is a technique for extracting from a given data set a dominant subspace that gives a good approximation of that data set. The method has been very successful for large data sets with a high level of redundancy [24, 42]. In simplified terms POD theory predicts [42] that if a parameter dependent vector $x(\Omega)$ depends smoothly on its parameter

$\Omega \in [\Omega_{min}, \Omega_{max}]$ and the vectors $x_1, \ldots, x_k$ are given as snapshots of x (i.e., $x_i = x(\Omega_i)$ for some nodes $\Omega_1, \ldots, \Omega_k \in [\Omega_{min}, \Omega_{max}]$), then $x(\Omega)$ can be well-approximated by a linear combination of $x_1, \ldots, x_k$ *for all values of* Ω. Thus, $x(\Omega)$ approximately lives in a small dimensional subspace spanned by $x_1, \ldots, x_k$. Moreover, a further reduction of the dimension is possible by selecting only the dominant directions in the family $x_1, \ldots, x_k$.

Coming back to the parameter dependent eigenvalue problem (2.2), we have that eigenvalue/eigenvector pairs, mode shapes, or, more general, invariant subspaces depend smoothly on the parameter (under the condition that they are simple which is generically the case)[37]. So, applying the POD methodology, in order to obtain a subspace $\mathcal{Q}$ associated with approximations to the right half plane eigenvalues for all Ω, it is enough to compute this subspace for some $\Omega_1, \ldots, \Omega_k \in [\Omega_{min}, \Omega_{max}]$, $i = 1, 2 \ldots k$ within the relevant region of Ω. Let $X(\Omega_i)$ denote an $n \times m_i$ matrix containing the unit-norm eigenvectors of P_{Ω_i} associated with the right half plane eigenvalues in its columns. Note that the number of columns m_i may change with Ω as the number of right half plane eigenvalues may change (it is piecewise constant in Ω). Since $X(\Omega_i)$ may contain complex numbers, whereas we look to have a real subspace, we use the real snapshot matrix

$$X = [\mathrm{Re}[X(\Omega_1), X(\Omega_2), ...X(\Omega_k)], \mathrm{Im}[X(\Omega_1), X(\Omega_2), ...X(\Omega_k)]] \in \mathbb{R}^{n,m} \tag{5.1}$$

with $m = 2 \sum_{i=1}^{k} m_i$.

Finally, truncating an (economy) SVD (see, e.g., [9]) of the measurement matrix X

$$X = [u_1, \ldots, u_m] \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_m \end{bmatrix} [v_1, \ldots, v_m]^T \tag{5.2}$$

is used to distill the projection space $\mathcal{Q}^{(j)}$ consisting of the left singular vectors corresponding to the singular values σ_i above a prespecified threshold s_ε, i.e., we choose $Q = [u_1, \ldots, u_d]$.

5.1. Implementation details

The POD approach of computing an approximation to the subspace associated with eigenvalues in a region $\mathcal{R} \subset \mathbb{C}$ for all $\Omega \in [\Omega_{\min}, \Omega_{\max}]$ is summarized in Algorithm 1.

Algorithm 1 POD for parametric QEPs

Input: $M, D_\Omega, K_\Omega, \Omega_{\min}, \Omega_{\max}, k$, region $\mathcal{R} \subset \mathbb{C}$, threshold s_ε
Output: Q
1: chose $\Omega_1, \ldots, \Omega_k \in [\Omega_{min}, \Omega_{max}]$
2: **for** $i = 1$ to k **do**
3:     Compute X_i as the set of eigenvectors (to eigenvalues in $\mathcal{R}$) of $(M, D(\Omega_i), K(\Omega_i))$
4: **end for**
5: $X = [\mathrm{Re}[X_1, X_2, \cdots, X_k], \mathrm{Im}[X_1, X_2, \cdots, X_k]]$
6: Compute an economy-size SVD $X = \sum_{i=1}^{m} \sigma_i^{(j)} u_i^{(j)} v_i^{(j)T}$
7: Set $Q = [u_1, u_2, \cdots u_d]$, with $\sigma_d \geq s_\varepsilon > \sigma_{d+1}$
8: **if** Q is not good enough **then** increase k and/or decrease s_ε **end if**

In the following we will discuss the important steps of the method, indicate the challenges and difficulties and describe the measures that we have taken in order to overcome these problems.

Input: The user has to specify the parametric QEP by providing: the FE matrices, M, D_Ω, K_Ω, the region $\mathcal{R}$ where the eigenvalues are sought, the parameter range $[\Omega_{\min}, \Omega_{\max}]$, a threshold s_ε for singular values, and the number of snapshots k.

line 1: Here a number of choices are possible to determine $\Omega_1, \ldots, \Omega_k$: uniform distribution and Chebychev points make sense [8].

line 3: The computing time of line 3 can be significantly reduced by using good starting values for the Krylov subspace method obtained by interpolating already computed eigenvectors.

line 6: If m is not too large, then X is a tall skinny matrix and then the economy-size SVD (5.2) can be efficiently computed by dense methods in $\mathcal{O}(nm^2)$ operations. Note that we do not need the complete SVD. So, if m is large, then it is more efficient to compute only the dominant singular vectors $u_1, \ldots, u_d$ by Krylov subspace methods requiring $\mathcal{O}(n(m + d^2))$ operations [9].

line 8: If the computed Q is not satisfactory (e.g., because the approximate eigenvectors obtained from it are not accurate enough or because an eigenvalue in the region $\mathcal{R}$ has been missed), then we can restart the algorithm with decreased tolerance s_ε (i.e., using an increased d) and/or increased k.

In a second run of the algorithm a lot of work can be saved by recycling results from the first run. For example, when d is increased, then X does not have to be recomputed. Furthermore, improvements can be obtained by updating the truncated singular value decomposition, again computed via a Krylov subspace method, in which the previously computed singular vectors can be recycled.

If k is increased, then this should be done by keeping the old node points and just adding new ones, e.g., by adding the mid points. For each new node point Ω_i, we append the eigenvectors corresponding to newly added parameter values as new columns to the measurement matrix X. Again, computing the SVD can be sped up by recycling the previous singular vectors.

6. Numerical experiments

In this section we compare the results of our new POD based model reduction approach with the traditional model reduction approach. As test problems, we use a simple academic brake model (model M0 with 4669 degrees of freedom) and two finite element models of industrial brakes: model M1 with around 1.2 million degrees of freedom and model M2 with $800,000$ degrees of freedom. These models are courtesy of our industrial partners.
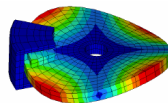
6.1. Academic brake model M0

Figure 6.1: An academic brake model M0 with 4669 dof

As a first example, and in order to validate the new model reduction approach, we use model DEVX6 [22, Section 8.3], a relatively coarse grain FE model of a brake with 4669 degrees of freedom (dof) as depicted in Fig. 6.1. This model uses a special damping term of the form $D_\Omega = D_1 + \frac{1}{\Omega}D_R + \Omega D_G + D_4/F_{ref}$, where D_4 accounts for frequency dependent damping. The coefficient matrices were constructed for a reference frequency $F_{ref} = 1600$Hz and a reference angular velocity $\Omega_{REF} = 1$ radians/sec. Some properties of the FE matrices are summarized in Table 1 and their sparsity patterns are plotted in Fig. 6.2. The range of variation of the parameter Ω is $[\Omega_{\min}, \Omega_{\max}] = [1, 4] \cdot 2\pi$.

We used Algorithm 1 to construct three POD models using $k = 3$, 5 or 9 snapshots at $\Omega_i = (1 + 3\frac{i-1}{k})2\pi$ for $i = 1, \ldots, k$. For the traditional approach we have set the maximum dimension of the subspace to 300, which is the default value of the model M2. We then used the generated reduced order models to compute approximate eigenvalues for $\Omega_{target} = 4\pi$ radians/sec within a rectangular domain $\mathcal{R} \subset \mathbb{C}$ given by $-50 < \text{Re}(\lambda) < 1000$ and $-1 < \text{Im}(\lambda) < 20,000$, corresponding to a frequency band from 0 Hz to 3 kHz.

The timings to reduce the model and then evaluate it for a value of Ω are given in Table 2. These experiments were carried out in MATLAB R2014b on a PC with a 4 core CPU (Intel Core i7 870 at 2.93GHz) and 8 GB of main memory. We observed that, when compared to the traditional method, the new method is more expensive in the reduction phase, but yields reduced models that are much more efficient to solve. The fact that the reduction phase is cheaper in the traditional

9

Table 1: Properties of the FE matrices for academic brake model M0. ("(skew-)symm" denotes (skew-)symmetry; rank denotes the structural rank; nnz denotes the number of non-zero entries)

matrix	pattern	2-norm	rank	nnz
M	symm	$6.09 \cdot 10^{-5}$	full	93 207
D_1	symm	6.86	337	5683
D_G	skew-symm	$1.52 \cdot 10^{-4}$	3564	66 643
D_R	symm	6.86	337	5683
D_4	symm	$3.22 \cdot 10^{4}$	419	23 885
K_1	symm	$5.98 \cdot 10^{7}$	full	280 845
K_R	neither	$1.45 \cdot 10^{5}$	337	21 646
K_{GEO}	symm	$2.40 \cdot 10^{-3}$	full	280 871

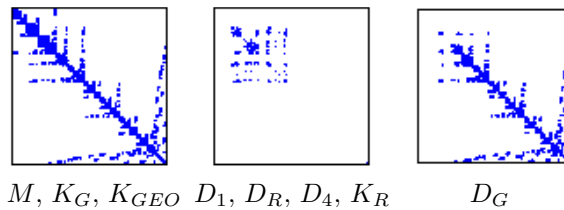M, K_G, K_{GEO} D_1, D_R, D_4, K_R D_G

Figure 6.2: Sparsity pattern for FE matrices associated with academic brake model M0

approach can be explained by the fact that the dominant operations during the reduction phase are matrix factorizations: the traditional approach performs one real Cholesky decomposition, whereas the new approach requires several complex LU decompositions. On the other hand the obtained reduced order models of the new method are typically of much smaller dimension and thus faster to solve. Furthermore, the size of the reduced models can be adapted to given accuracy requirements, since an error estimate from the singular value decomposition is available [9].

In order to assess the quality of the reduced order models, we measure the distance between the eigenvalue/eigenvector pairs $(\lambda_i, x_i)_{i=1,2,\dots}$ in the region $\mathcal{R}$ of the original system and the approximated ones $(\lambda_j^r, x_j^r)_{j=1,2,\dots}$ of the reduced systems in three different ways. The first is the maximal distance of a true eigenvalue to the nearest approximated one, relative to the modulus of the largest computed eigenvalue.

$$err_\lambda := \max_i(\min_j |\lambda_i - \lambda_j^r|)/\max_i(|\lambda_i|). \tag{6.1}$$

To analyze the accuracy of the approximated eigenvectors, we use the largest acute angles between an exact eigenvector and the closest approximation, and between an exact eigenvector and the space $\mathcal{Q}$ spanned by the columns of Q,

$$err_x := \max_i(\min_j \angle(x_i, Qx_j^r)), \qquad\qquad err_Q := \max_i \angle(x_i, \mathcal{Q}), \tag{6.2}$$

Table 2: Computation times for the M0 model

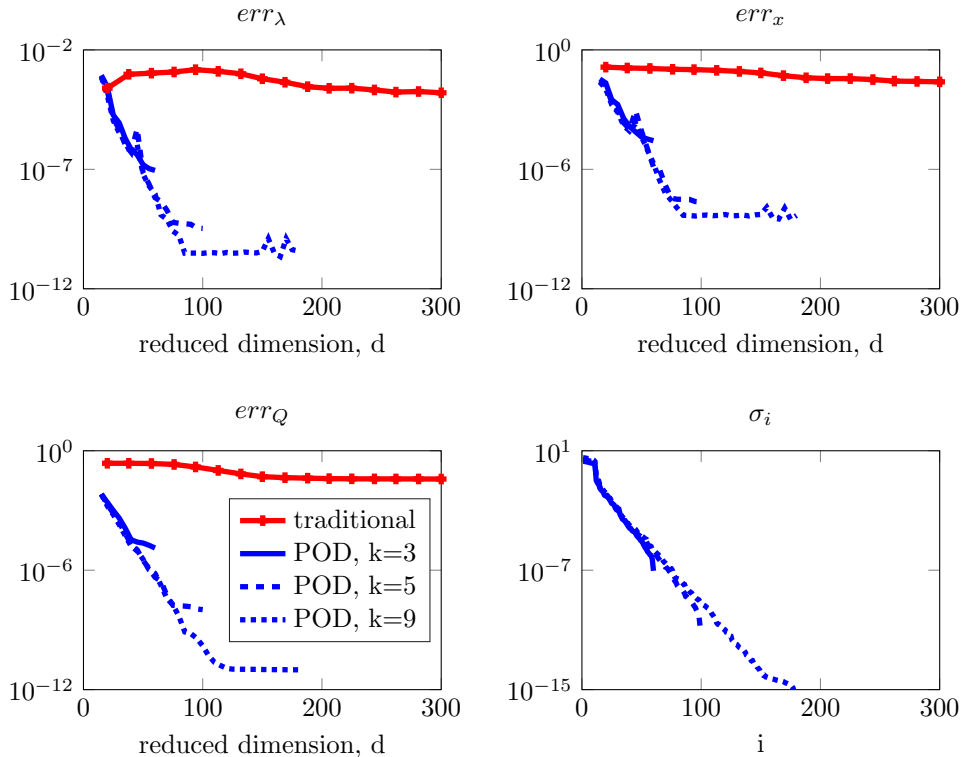|	reduction	evaluation
without reduction	0 s	7.1 s
traditional, $d = 300$	8.7 s	2.3 s
POD, $k = 3$, $d = 40$	10.5 s	0.016 s
POD, $k = 5$, $d = 75$	17.3 s	0.07 s
POD, $k = 9$, $d = 85$	31.5 s	0.09 s

10

Figure 6.3: Relation of the various error metrics and the reduced dimension for the considered Algorithms for model M0.

respectively. A low value of err_Q is necessary, but not sufficient for a low value of err_x.

The results are plotted in Fig. 6.3. It can be seen from these figures that the POD based approach is more accurate, and at the same time yields a smaller dimension of the projection subspace than the traditional subspace.

Let us mention some noteworthy details: The traditional approach does not benefit much from increasing d: for example, driving d from 20 to 300 improves err_x and err_Q only by one order of magnitude. Moreover, err_λ even increases in the early stages and only for $d = 200$ it reaches the same value as for $d = 20$. The POD approach on the other hand benefits heavily from increasing d. Roughly speaking, increasing d by 10 results in an order of magnitude better accuracy. And for small d (till ≈ 40) the accuracy does not depend on the number of snapshots. After that increasing d for $k = 3$ is not as effective as for $k = 5$, or 9. For $k = 5$ the sweat spot seems to be at $d = 75$, and for $k = 9$ at $d = 85$. This is most apparent form err_Q (which, in contrast to err_λ and err_x, is by definition monotonically decreasing). The figures clearly indicate that the POD approach can be used for an adaptive subspace selection.

In the next experiment we fixed d (to 40 for $k = 3$, to 75 for $k = 5$, and to 85 for $k = 9$) and varied Ω. The results are plotted in Fig. 6.4. It can be observed that the POD approach is consistently more accurate for both eigenvalues and eigenvectors for the whole range of Ω.

6.2. Results for industrial brake model M1

Next we tested the performance of the POD method by applying it to a brake model (called M1) from an industrial partner. This model has over 1.2 million degrees of freedom. The properties of the FE matrices are summarized in Table 3. The matrices were constructed for $\Omega_{REF} = 1$, $F_{ref} = 3$ kHz. The parameter Ω varies within $[2\pi, 8\pi]$ and we are interested in the eigenvalues at $\Omega_{target} = 4\pi$ within a rectangular domain $\mathcal{R}$ given by $-1 < \mathrm{Re}(\lambda) < 1000$ and $-1 < \mathrm{Im}(\lambda) < 20,000$. The imaginary part corresponds to a frequency band from 1 kHz to 3 kHz. We used
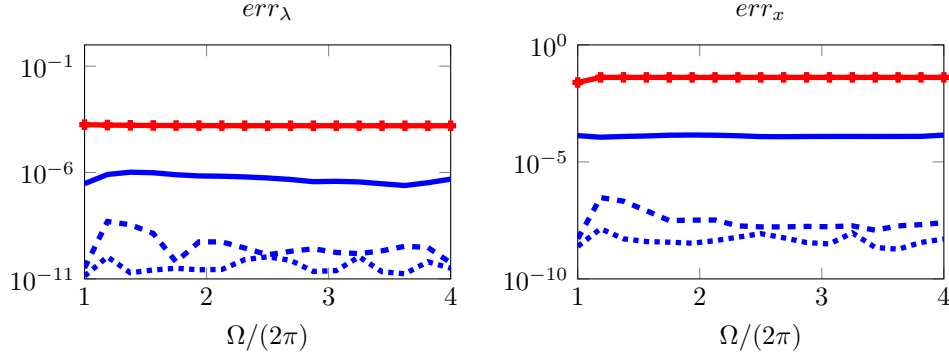
Figure 6.4: Model M0: Accuracy of the computed eigenvalues (left) and eigenvectors (right) plotted over Ω. The legend is the same as in Fig. 6.3

Table 3: Properties of FE matrices of industrial brake model M1

matrix	pattern	2-norm	rank	nnz
M	symm	3.2e-2	Full	34 826 213
D_1	symm	6.1e-1	3833	80 567
D_G	skew-symm	1.4e-4	180 991	3 533 431
D_R	symm	1.1e-1	3820	80 542
D_4	symm	1.5e9	Full	95 927 033
K_1	symm	3.1e11	Full	95 928 131
K_R	neither	1.1e5	3820	491 405
K_{GEO}	symm	1.8e1	Full	47 068 113

$k = 3, 5, 9$ snapshots and a threshold $s_\varepsilon = 1e-12$ for truncating singular values. The convergence behavior of the quantities of interest (eigenvalues and eigenvectors) are shown in Fig. 6.5, where the error metrics $err_\lambda, err_x, err_Q$ as defined in (6.1), and (6.2) are used. We observe that for $k = 3, 5,$ and 9 the POD approach is less, equally, and way more accurate than the traditional approach, respectively. These experiments were carried out in MATLAB R2013a on a PC with 8 cores (dual Intel Xeon E5-2643 CPU at 3.3GHz) and 128GB main memory.

We note that this brake model is very ill-conditioned (more details on ill-conditioning and sensitivity of eigenvalues are given in Section 7 below). So, even the best approximations of eigenvalues denoted by ("exact"), i.e., those computed by Alg. 2 for $\Omega = \Omega_{target}$) cannot be fully trusted. Thus, the results of Fig. 6.6 should be considered with a bit of care, and hence, we additionally calculated the residual of selected computed eigenvalues via

$$r(\lambda, x) := \frac{\|(\lambda^2 M + \lambda C + K)x\|_\infty}{\|(|\lambda|^2|M| + |\lambda||C| + |K|)|x|\|_\infty},$$

where $|M|$ denotes the matrix of absolute values of the elements of M.

Fig. 6.6(left) is typical for the results we obtained. We see two eigenvalues at roughly $1,800$ and 2800Hz, both computed by Algorithm 2 ("exact") as well as by model reduction with the traditional and the POD approach. The exact and POD computed eigenvalues lie on top of each other, the traditional ones are close but noticeably distinct. Moreover, the residuals r (indicated by the gray level of the shown marks (o,+,x) in Fig. 6.6) of the eigenvalues computed by the traditional method ($\approx 10^{-4}$) is way higher than those of the "exact" and POD ones ($\approx 10^{-11}$).

6.3. Results for industrial brake model M2

In this section we discuss the FE matrices coming from a second industrial brake model M2. This model is shown in Fig. 2.1. It uses a damping term given by $D = D_1 + \frac{1}{\Omega}D_R + \Omega D_G$. The FE matrices for this model have around $800,000$ dof. We built the POD measurement matrix of the
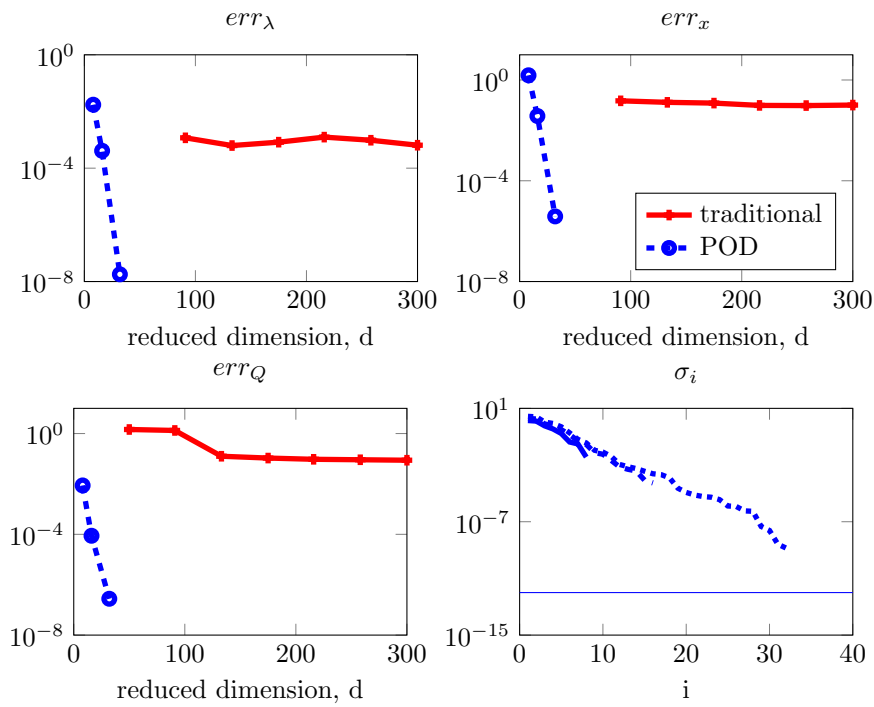
Figure 6.5: Relation of the various error metrics and the reduced dimension for the considered Algorithms for model M1.
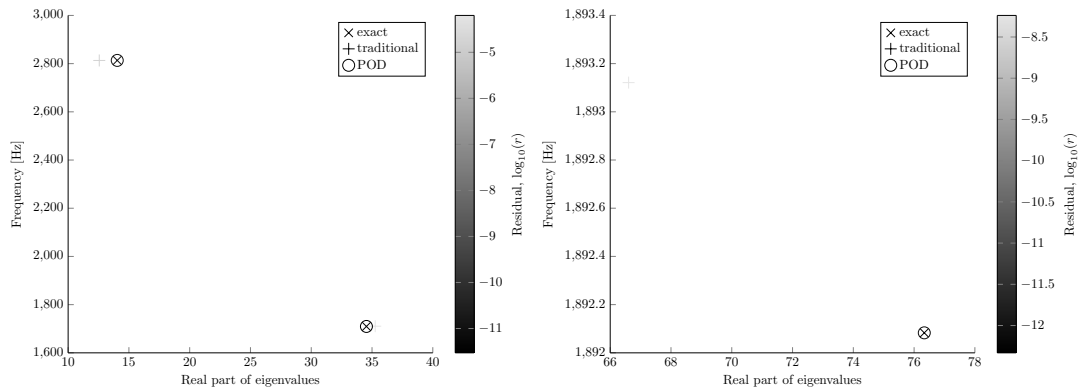

Figure 6.6: Industrial models M1(left) and M2(right): Selected eigenvalues for different methods: POD, traditional and full calculation, color coded with their residuals
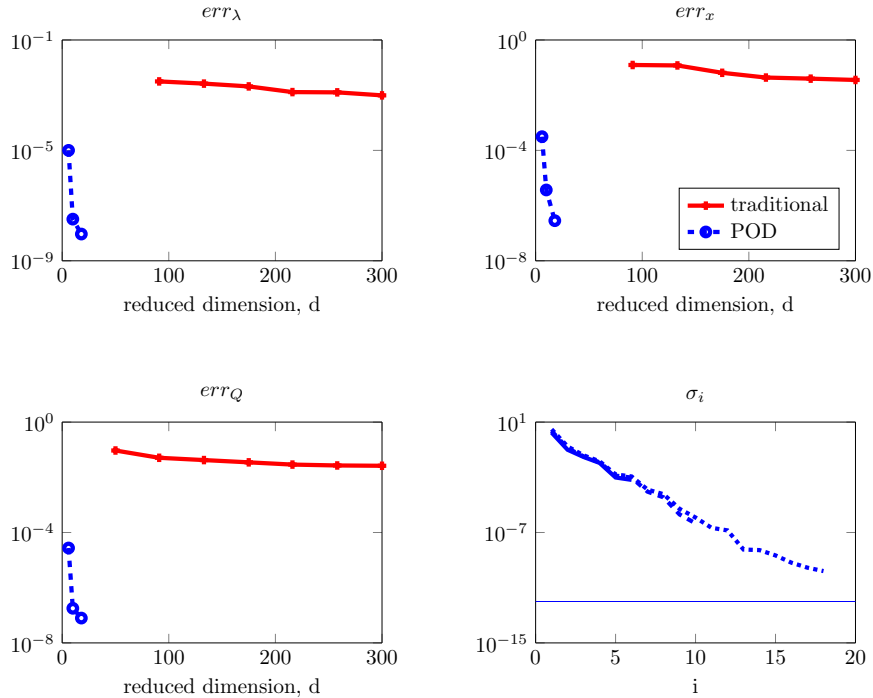
Figure 6.7: Relation of the various error metrics and the reduced dimension for the considered Algorithms for model M2.

QEP (5.1) by solving the QEP for parameter values in $\Omega = [2\pi, 8\pi]$, and measure the results for a parameter value $\Omega_{target} = 4\pi$. The convergence behavior and the eigenvalues obtained from POD and the traditional approaches are depicted in Fig. 6.7 and Fig. 6.6(right), respectively. We note that the conditioning of this model is even worse than that model M1. The results are similar to those of models M0 and M1: the POD approach is computationally more expensive, but leads to smaller reduced order models whose eigenvalue approximations are way more accurate.

7. Eigenvalue sensitivity

The eigenvalues associated with the disk brake model are the values λ such that the matrix $P_\Omega(\lambda)$ is not invertible. Usually, there are $2n$ eigenvalues and all $2n$ eigenvalues are finite if the mass matrix M is invertible. If M is singular, which happens if rigid connections are modeled as constraints, then there may be infinite eigenvalues (provided that $P_\Omega(\lambda)$ is invertible for at least one value of λ).

However, if the matrix $P_\Omega(\lambda)$ is not invertible for any λ, then "all complex numbers are eigenvalues" and the eigenvalue problem is called singular. This situation usually does not make sense physically and is a clear indicator of a mistake in the FE modeling. In our tests with the industrial brake models we did not have a singular eigenvalue problem, but in some cases the matrix $P_\Omega(\lambda)$ was close to a non-invertible matrix for all λ and Ω. In such a case the eigenvalues and eigenvectors are extremely sensitive to small perturbations. The computed eigenvalues and eigenvectors are typically corrupted with large errors and cannot be trusted. Thus, this property needs to be tested and detected in the course of a numerical method, because it again is an indicator of bad FE modeling.

One way to quantify the sensitivity of an eigenvalue problem is via its pseudospectrum and backward errors using perturbation analysis. The backward error Δ_b of an approximate eigenvalue $\tilde{\lambda}$ is the smallest possible norm of a perturbation to the original matrices M, D, K, making $\tilde{\lambda}$ an
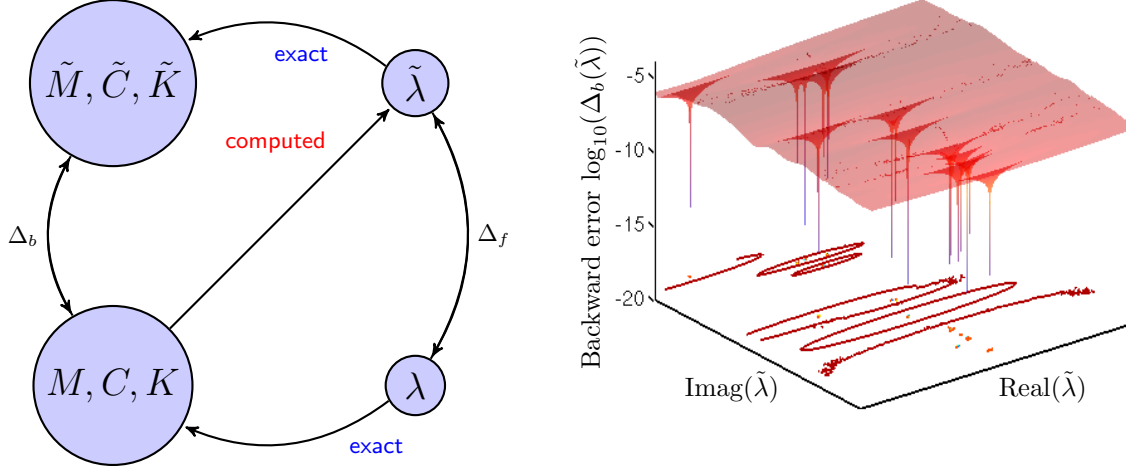
Figure 7.1: (left) Relation between forward error Δ_f and backward error Δ_b. (right) Surface plot of the backward error Δ_b at various points of the complex plane $\tilde{\lambda}$

exact eigenvalue of the resulting nearby matrices $\widetilde{M}, \widetilde{D}, \widetilde{K}$:

$$\Delta_b(\tilde{\lambda}) := \min \left\{ \varepsilon > 0 : \exists \widetilde{M}, \widetilde{D}, \widetilde{K}, \tilde{x} \text{ such that} \right.$$
$$\left. \|\widetilde{M} - M\| < \varepsilon \|M\|, \|\widetilde{D} - D\| < \varepsilon \|D\|, \|\widetilde{K} - K\| < \varepsilon \|K\|, (\tilde{\lambda}^2 \widetilde{M} + \tilde{\lambda} \widetilde{D} + \widetilde{K}) \tilde{x} = 0 \right\}.$$

Fig. 7.1(left) shows the relationship between data matrices, eigenvalues, forward error and backward error. The value Δ_b can be computed for any $\tilde{\lambda}$ in the complex plane by [17]

$$\Delta_b(\tilde{\lambda}) = \frac{\sigma_{\min}(\tilde{\lambda}^2 M + \tilde{\lambda} D + K)}{|\tilde{\lambda}|^2 \|M\| + |\tilde{\lambda}| \|D\| + \|K\|}, \tag{7.1}$$

where $\sigma_{\min}(\cdot)$ denotes the smallest singular value of a matrix.

We have that $\Delta_b(\tilde{\lambda}) \geq 0$ for all $\tilde{\lambda}$ and it is zero only for eigenvalues. Numerically, a value $\tilde{\lambda}$ is considered an eigenvalue if its backward error $\Delta_b(\tilde{\lambda})$ is below a small tolerance $\epsilon \ll 1$, usually on the level of the machine precision. Each value of ϵ defines a set of approximate eigenvalues, $\Lambda_\epsilon := \{\tilde{\lambda} \in \mathbb{C} : \Delta_b(\tilde{\lambda}) \leq \epsilon\}$, called the ϵ-pseudospectrum [39]. For a numerically well-behaved model the pseudospectrum contains the exact eigenvalues and a very small area around them. For an example see Fig. 7.1(right) where the backward error of the model M0 (see Section 6) is depicted in a surface plot. Additionally, contour lines and the position of the eigenvalues are marked in the bottom. We see that Δ_b is relatively large ($\approx 10^{-6}$) almost everywhere with steep downward peaks reaching machine precision tightly around the eigenvalues. So any value $\tilde{\lambda}$ with small backward error is close to an exact eigenvalue.

If, however, the pseudospectrum covers a large region for small values of ϵ, then the eigenvalues cannot be precisely calculated in finite precision arithmetic, and the eigenvalue problem is called numerically singular. Section Appendix B suggests a few practical tricks to remodel an FE system, which reduces the sensitivity of eigenvalues in response to perturbations in the data (FE matrices). Let us consider the two industrial models M1 and M2. Since evaluating formula (7.1) is costly for large n, creating a plot like 7.1(right) is out of reach for these models. Instead we computed the backward error for 20 values of $\tilde{\lambda}$ randomly chosen in the region $\mathcal{R}$ of interest. We found Δ_b to be about 10^{-11} and 10^{-14}, respectively, (and thus very close to the machine precision) for all tested values of $\tilde{\lambda}$. Since it is extremely unlikely that all chosen values were close to eigenvalues, we consider the model M2 to be numerically singular (and to a lesser extend also model M1), and we recommend to remodel the brake in both cases.

8. Conclusion

In this paper we have presented an approach using proper orthogonal decomposition (POD) for model reduction of dynamical systems in the context of a brake squeal problem. The POD approach is based on the solution of a parametric quadratic eigenvalue problem, taking into account all the important effects including damping and friction induced terms. In situations where these parameters play a significant role, the POD method will have a clear edge over existing methods in terms of accuracy of the computed solution. For simpler problems, where the parametric effects, damping, friction and associated nonlinearities are not as significant, one can still make design decisions based on the traditional approach of modal truncation. In any case it is recommendable to check the sensitivity of the computed eigenvalues.

The open source Python program that resulted from this project will be useful not only in brake squeal studies, but also in other parameter dependent vibration problems which have non-proportional damping and other non-symmetric terms. We indicate how to detect numerical problems like high sensitivity associated with the FE modeling, and how to improve the solution of the eigenvalue problems via the scaling of the matrices.

With this method we have studied academic models as well as real industrial brake models. Our implementation can identify FE modeling problems which could lead to numerical inaccuracies. Once the FE model is corrected, either the traditional method or the POD method can be applied, depending on the desired accuracy of the solution.

Appendix A. Numerical solution of quadratic eigenvalue problems

There are many numerical methods for quadratic eigenvalue problems (QEPs). They all work for a fixed matrix triple (M, D, K). So, in a parameter study for a parametric QEP the following computational tasks have to be repeated for many values of the parameter Ω. The choice of the right numerical method to solve a QEP (3.1) depends largely on the problem dimension n.

For small scale problems (n up to several thousands) dense methods like the QZ method [9], are certainly a good choice. For large dimensions iterative projection methods should be used [2, 27]. These methods are meant to compute some exterior eigenvalues of a matrix whereas we need the eigenvalues in a region $\mathcal{R}$ of the right complex half plane. Thus, some preprocessing is necessary, such as spectral transformations [5]. In particular, we use the shift-scale-invert-linearize scheme, whose steps are described in the following.

shift: Choosing a shift value τ such that $\det(P_\Omega(\tau)) \neq 0$, we convert (3.1) to a new QEP of the form

$$((\lambda - \tau)^2 M_\tau + (\lambda - \tau)D_\tau + K_\tau)x = 0,$$

where $M_\tau = M$, $D_\tau = 2\tau M + D$ and $K_\tau = \tau^2 M + \tau D + K$ is nonsingular. The shift point τ is chosen in the positive half plane, ideally near the region where the eigenvalues are expected to be. These regions can be estimated, e.g., via Gersgorin-type theorems [40]. On how to choose the shift points τ for the spectral transformations, see Section Appendix A.1.

scale: The next step is an appropriate scaling of the FE matrices, since often the elements of M_τ and K_τ vary by large orders of magnitude. It is a well-established fact that computing with such matrices will result in a large loss of accuracy and stability of the computed eigenvalues and eigenvectors, see [17]. We scale the problem using the scalar scaling procedure in [6] and diagonal scaling before performing the linearization. For this we compute diagonal matrices $\mathcal{D}_L$ and $\mathcal{D}_R$ such that every row and every column of the matrix $\mathcal{D}_L Y \mathcal{D}_R$ is of unit 2-norm where

$$Y = |\gamma^2 \delta M_\tau| + |\gamma \delta D_\tau| + |\delta K_\tau|, \qquad \gamma = \sqrt{\|K_\tau\|/\|M_\tau\|}, \qquad \delta = 2/(\|K_\tau\| + \|D_\tau\|\gamma), \qquad \text{(A.1)}$$

and $\|\cdot\|$ denotes some efficiently computable matrix norm. The scaling is performed via $\widehat{M} = \gamma^2 \delta \mathcal{D}_L M_\tau \mathcal{D}_R$, $\widehat{D} = \gamma \delta \mathcal{D}_L D_\tau \mathcal{D}_R$, and $\widehat{K} = \delta \mathcal{D}_L K_\tau \mathcal{D}_R$, and the QEP is now of the form

$$((\tfrac{\lambda-\tau}{\gamma})^2 \widehat{M}_\tau + \tfrac{\lambda-\tau}{\gamma}\widehat{D}_\tau + \widehat{K}_\tau)\hat{x} = 0,$$

16

with $\hat{x} = \mathcal{D}_R^{-1}$.

invert: If we then consider the reverse polynomial, which just reverts the order of the coefficients and corresponds to replacing the (shifted and scaled) eigenvalue by its inverse $\mu_\tau = \frac{\gamma}{\lambda - \tau}$, then the eigenvalues near τ become large in modulus, while the ones far away from τ become small. Since eigenvalue methods typically converge to the exterior eigenvalues at first, we expect that this leads to fast convergence to the desired eigenvalues. We also (formally) invert $\widehat{K}_\tau$ to obtain a monic QEP

$$(\mu_\tau^2 I + \mu_\tau \widehat{K}_\tau^{-1} \widehat{D}_\tau + \widehat{K}_\tau^{-1} \widehat{M}_\tau)\hat{x} = 0, \tag{A.2}$$

linearize: We use the classical companion linearization (3.2) to pose (A.2) as a linear eigenvalue problem

$$\mathcal{A}_\tau y = \mu_\tau y, \quad \mathcal{A}_\tau = \begin{bmatrix} -\widehat{K}_\tau^{-1} \widehat{D}_\tau & -\widehat{K}_\tau^{-1} \widehat{M}_\tau \\ I_n & 0 \end{bmatrix} \tag{A.3}$$

where $y = [\mu_\tau \hat{x}^T, \hat{x}^T]^T$. The EVP (A.3) can be solved by the Arnoldi method which is implemented e.g., in `ARPACK` [27], `MATLAB`'s `eigs`, or `scipy.sparse.linalg.eigs` in Python. All that these methods require, is to compute matrix-vector products of the form $[y_1^T, y_2^T]^T = \mathcal{A}_\tau [x_1^T, x_2^T]^T$ for several given vectors $[x_1^T, x_2^T]^T$. These can be formed as

$$y_1 = -\widehat{K}_\tau^{-1}(\widehat{D}_\tau x_1 + \widehat{M}_\tau x_2), \quad y_2 = x_1. \tag{A.4}$$

Note that it is not necessary to form the inverse of $\widehat{K}_\tau$, but only its action on a vector which can be computed much cheaper by a (sparse) factorization (done once during initialization), or another iterative method.

Remark Appendix A.1. It is good practice to check $\widehat{K}_\tau$ for its condition number $\kappa(\widehat{K}_\tau) = \|\widehat{K}_\tau\| \|\widehat{K}_\tau^{-1}\|$ which can be cheaply approximated with the just computed matrix factorization [18, chapter 15]. When the condition number is very large (say $\geq 10^{12}$), then numerical difficulties are inevitable. This typically happens for two possible reasons, the first being that the shift τ is too close to an actual eigenvalue. In this case the Arnoldi method will likely not be able to compute an eigenvalue other than the one at τ. Choosing another shift usually cures this situation. The other possible reason for large $\kappa(\widehat{K}_\tau)$ is that the QEP is close to being singular. In this case, that can be recognized by a large $\kappa(\widehat{K}_\tau)$ for several values of τ, the eigenvalues λ are not well-defined, in the sense that tiny perturbations of the matrices M, D, K can completely change the location of the eigenvalues. It is best to return back to the modeling stage, since this indicates a bad FE model. If possible, the finite element nodes that are responsible for the high condition number in K_τ are identified using extremal eigenvectors, via the procedure described in Section Appendix B.

Remark Appendix A.2. As described, we use the shift-invert spectral transformation $\mu_\tau = \gamma/(\lambda - \tau)$. With Krylov methods tending to converge towards the eigenvalues μ_τ that are large in modulus first, we expect rapid convergence towards eigenvalues λ close to τ. We may also consider the alternative spectral transformation $\theta_\tau := (\lambda + \bar{\tau})/(\lambda - \tau)$ which has the favorable property that $\text{Re}(\lambda) \geq 0 \Leftrightarrow |\theta_\tau| \geq 1$ for $\text{Re}(\tau) > 0$. So, a Krylov method using this transformation converges towards the eigenvalues in the right half plane. It turns out that we are implicitly using this transformation. This is because $\theta = (\tau + \bar{\tau})\mu + 1$ is a scaled and shifted value of μ and Krylov methods are invariant under scaling and shifting.

Once eigenvalues/eigenvectors $(\mu_\tau, [y_1^T, y_2^T]^T)$ of $\mathcal{A}_\tau$ have been computed, we obtain eigenvalues/eigenvectors (λ, x) of the original problem (M, D, K) by undoing the spectral transformations:

$$\lambda = \tau + \gamma/\mu_\tau, \qquad x = \begin{cases} \mathcal{D}_R y_1/\|\mathcal{D}_R y_1\|_2, & |\mu_\tau| \geq 1, \\ \mathcal{D}_R y_2/\|\mathcal{D}_R y_2\|_2, & \text{else.} \end{cases} \tag{A.5}$$

The overall process is summarized in Algorithm 2. The dominant computational cost are usually the LU factorizations, which have to be done for every shift τ. So, a parameter study with many values of Ω requires the solution of a large number of large scale eigenvalue problems.

Algorithm 2 Numerical solution of large QEPs

Input: M, D, K

1: **repeat**
2:     choose target τ in right half plane, see Sec Appendix A.1
3:     set $K_\tau := \tau^2 M + \tau D + K$, $D_\tau := 2\tau M + D$, $M_\tau := M$
4:     compute $\gamma, \delta, \mathcal{D}_L, \mathcal{D}_R$ by (A.1)
5:     set $\widehat{M}_\tau = \gamma^2 \delta \mathcal{D}_L M_\tau \mathcal{D}_R$, $\widehat{D}_\tau = \gamma \delta \mathcal{D}_L D_\tau \mathcal{D}_R$, $\widehat{K}_\tau = \delta \mathcal{D}_L K_\tau \mathcal{D}_R$
6:     compute LU decomposition of $\widehat{K}_\tau$
7:     **if** $\kappa(K_\tau)$ too large **then** signal numerical problem **end if**
8:     apply a Krylov method $\mathcal{A}_\tau$ using (A.4)
9:     **for** every found eigenvalue/eigenvector $(\mu_\tau, [y_1^T, y_2^T]^T)$ of $\mathcal{A}_\tau$ **do**
10:         get (λ, x) by (A.5)
11:     **end for**
12: **until** all eigenvalues in $\mathcal{R}$ are found

Algorithm 3 Random sampling algorithm for choosing shift points

---

Input: A prespecified region $\mathcal{R}$.
Output: Shift points: τ_i, $i = 0, \ldots, m$

1: Set τ_0 as the center of gravity of the domain $\mathcal{R}$
2: $m = 1$
3: Call the Arnoldi method to get the radius r_0 of the covered circle at shift τ_0 and compute the area fraction (a_c) covered by this circle
4: **while** $a_c < 1$ **do**
5:     randomly throw 100 circles into the region $\mathcal{R}$ outside the covered area and with random radius with mean $r_{m,predict} = \frac{1}{m} \sum_{i=1}^{m} r_i$
6:     compute a_c and choose the center τ_m which resulted in highest area fraction a_c
7:     Call Arnoldi method to get the radius $r_{m,actual}$ of the covered circle at shift τ_m and compute the fraction of area a_c covered by this circle
8:     $m = m + 1$
9: **end while**

Appendix A.1. Random shift point selection

We have designed a heuristic random sampling algorithm to choose shift points for the shift-and-invert Arnoldi method. The idea is to cover the domain $\mathcal{R}$ completely using circles (centered at the shift points) using a minimum number of shifts. This avoids unnecessary LU-factorizations of K_τ and calls to `ARPACK`.

The algorithm is described in Algorithm 3. We randomly choose several trial shifts and randomly assign an expected radius to each shift. The algorithm then picks the best location for placing a shift point so that it covers a maximum area. This new location can be the largest uncovered region of the domain. For an initial uncovered convex domain, this is usually the center of gravity of the domain. In order to determine the location for subsequent shifts, we look at random locations in the uncovered plane by throwing 100 circles with a radius that is Gaussian distributed around the mean of radii from previous shifts. We compute the area fraction again using Monte Carlo integration, covered by these random shifts and then pick the one shift which results in the greatest area fraction from these 100 trials. We set this shift location as the parameter in the Arnoldi method and compute the eigenvalues around it and continue until the whole domain is covered completely.

Appendix B. Remodeling of ill-conditioned systems

It has been reported recently in [23], that almost-singularity of the QEP happens frequently for FE models as a result of poor judgment and bad modeling choices. It was shown there that the
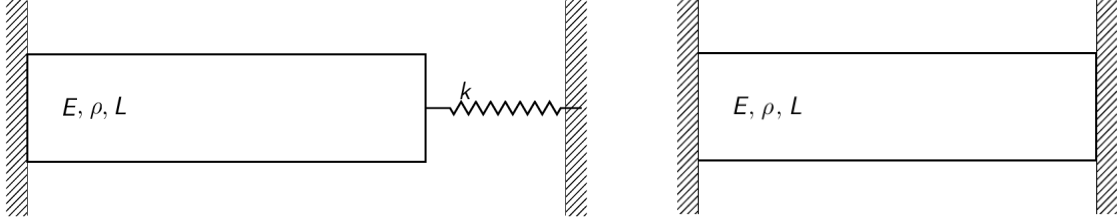
Figure B.1: left: A system with disproportionate stiffness, right: Remodeled FE system, parameters for both: $E = 2 \cdot 10^{11} \mathrm{N/m^2}$, $\rho = 7800 \mathrm{kg/m^3}$, $L = 1\mathrm{m}$, $k = 10^{23} \mathrm{N/m}$

modeling choices that lead to almost singularity can be traced by studying the eigenvectors corresponding to extremal eigenvalues of the stiffness matrix K. Moreover, there is a relation between the almost singularity of a quadratic eigenvalue problem arising from FE modeling and disproportionately large values of relative virtual energies of individual finite elements. Let us call the eigenvector associated to the smallest eigenvalue the 'smallest eigenvector' and the eigenvector associated to the largest eigenvalue the 'largest eigenvector'. The elements causing almost-singularity are those with high virtual kinetic energy for the smallest eigenvectors, or of high virtual strain energy for the largest eigenvectors. Inspecting those elements, one can identify problematic elements, and the user can change the FE model in order to improve the conditioning of the stiffness matrix K.

A similar effect also occurs if a rigid connection is replaced by a very stiff spring, a modeling trick often used in industrial practice. We illustrate the remodeling procedure by a simple example. Consider an axial bar which is connected to a very stiff spring with stiffness $k_0 = 10^{23}$ as shown in Fig. B.1(left). The equation describing the axial vibration of a bar can be written as

$$\frac{\partial}{\partial x}(EA\frac{\partial u}{\partial x}) = \rho A \frac{\partial^2 u}{\partial t^2},$$

where, E is Young's modulus of elasticity, A is the cross section area of the bar, ρ is the material density. By a suitable ansatz the time dependent terms can be eliminated, resulting in the following differential equation

$$\frac{d}{dx}(EA\frac{du}{dx}) + \rho A u = 0.$$

Assuming $E = 2 \times 10^{11}$, $\rho = 7800$, and $L = 1$, we discretize the eigenvalue problem using linear finite elements and obtain the generalized eigenvalue problem

$$(\lambda M + K)u = 0,$$

where K and M are the stiffness and mass matrices, the eigenvalues describe vibration frequencies, and the eigenvectors describes the mode shapes. The high stiffness of the spring on the right end of the bar results in eigenvalues that are highly sensitive to perturbations in the matrices, see Section 7. We calculate the backward error Δ_b by (7.1) to measure the sensitivity of eigenvalues to perturbations in the data. $\Delta_b(\tilde{\lambda})$ is plotted for a large number of real values of $\tilde{\lambda}$ in Fig. B.2(left). Since $\Delta_b(\tilde{\lambda})$ is below machine precision ($\approx 10^{-16}$) for the whole range of $\tilde{\lambda}$, we cannot localize the eigenvalues, even if we assume the matrices (which depend on parameters such as E, ρ) are correct to machine precision (16 decimal places).

Following [23], we can determine the parts of the FE model which need remodeling. The largest eigenvector of K is approximately sparse and has a dominant entry at the node where the high stiffness spring is connected. The user can remodel the system by replacing the stiff spring by a rigid connection, in terms of matrices this corresponds to the removal of the rows and columns corresponding to the high stiffness spring. The modified structure is depicted in Fig. B.1(right) and the backward error of the corresponding FE model is plotted in Fig. B.2(right). With this remodeled system the eigenvalues can be located to good precision, if we assume the matrices
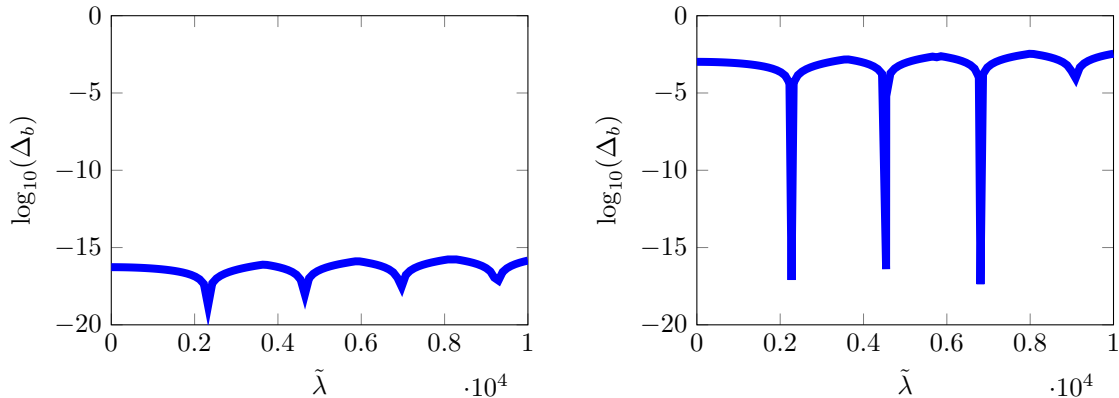
19

Figure B.2: Backward error of a system with disproportionately high stiffness (left), and of the remodeled system (right)

are accurate to 5 decimal places (i.e., assuming that material constants E, ρ, etc. which are used to construct matrices can be measured with accuracy of 5 significant places). The material properties used in brake squeal simulation involve many other material constants like friction coefficient, damping, and the backward error depends on the accuracy of these input data. It is dangerous to assume machine precision accuracy when the precision in not specified a priori, as this could lead to erroneous results, since the accuracy of the results depends on the accuracy of the matrices.

In view of the negative effect on the sensitivity of the eigenvalue problems, the replacement of rigid connections by stiff springs is not recommended.

Acknowledgments

References

[1] A. Akay. Acoustics of friction. *Acoustical Society of America Journal*, 111:1525–1548, Apr. 2002.

[2] Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst, editors. *Templates for the Solution of Algebraic Eigenvalue Problems*. Software, Environments and Tools. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2000.

[3] S. L. Campbell. Linearization of DAE's along trajectories. *Zeitschrift für angewandte Mathematik und Physik*, 46:70–84, 1995.

[4] D. Dowson. *History of Tribology*. Professional Engineering Publishing, Bury St Edmunds, Suffolk, 1998.

[5] T. Ericsson and A. Ruhe. The spectral transformation lanczos method for the numerical solution of large sparse generalized symmetric eigenvalue problems. *Mathematics of Computation*, 35:1251–1268, 1980.

[6] H. Fan, W. Lin, and P. V. Dooren. Normwise scaling of second order polynomial matrices. *SIAM J. Matrix Anal. Appl.*, 26,1:252–256, 2004.

[7] I. Gohberg, P. Lancaster, and L. Rodman. *Matrix Polynomials*. Academic Press, New York, 1982.

[8] G. Golub and G. Meurant. *Matrices, Moments and Quadrature with Applications*. Princeton University Press, Princeton, N.J., 2009.

[9] G. Golub and C. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore, Maryland, 3rd edition, 1996.

[10] N. Gräbner, S. M. Quraishi, C. Schröder, V. Mehrmann, and U. von Wagner. New numerical methods for the complex eigenvalue analysis of disk brake squeal. In *Proceedings of Eurobrake Conference, Lille, France*, 2014.

[11] N. Gräbner, M. Tiedemann, U. Von Wagner, and N. Hoffmann. Nonlinearities in friction brake NVH - experimental and numerical studies. Technical Paper 2014-01-2511, SAE, 2014.

[12] A. Guran, F. Pfeiffer, and K. Popp, editors. *Dynamics with Friction: Modeling, Analysis and Experiment*, volume 7 of *Series on stability, vibration and control of systems*. World Scientific Pub., 1996. Part I.

[13] A. Guran, F. Pfeiffer, and K. Popp, editors. *Dynamics with Friction: Modeling, Analysis and Experiment*, volume 7 of *Series on stability, vibration and control of systems*. World Scientific Pub., 2001. Part II.

[14] W. Hahn. *Stability of Motion*. Springer-Verlag, Berlin, Germany, 1967.

[15] S. Hammarling, C. J. Munro, and F. Tisseur. An algorithm for the complete solution of quadratic eigenvalue problems. *ACM Trans. Math. Softw.*, 39(3):18:1–18:19, May 2013.

[16] P. Hartman. A lemma in the theory of structural stability of differential equations. *Proceedings of the American Mathematical Society*, 11(4):pp. 610–620, 1960.

[17] N. Higham, D. Mackey, F. Tisseur, and S. Garvey. Scaling, sensitivity and stability in the numerical solution of quadratic eigenvalue problems. *International Journal of Numerical Methods in Engineering*, 73:344–360, 2008.

[18] N. J. Higham. *Accuracy and stability of numerical algorithms*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second edition, 2002.

[19] N. J. Higham, D. S. Mackey, N. Mackey, and F. Tisseur. Symmetric linearizations for matrix polynomials. *SIAM J. Matrix Anal. Appl.*, 29(1):143–159 (electronic), 2006/07.

[20] D. Hochlenert and U. von Wagner. How do nonlinearities influence brake squeal? In *Proceedings of 29th SAE brake colloquium, pp. 179-186*, 2011.

[21] INTES, Stuttgart. *PERMAS Product Description*, 2012. Version 14.

[22] INTES, Stuttgart. *PERMAS Example Manual*, 2014. Version 14.00.147, INTES publication no. 550.

[23] R. Kannan, S. Hendry, N. J. Higham, and F. Tisseur. Detecting the causes of ill-conditioning in structural finite element models. *Comput. Struct.*, 133:79–89, Mar. 2014.

[24] G. Kerschen, J. Golinval, A. Vakakis, and L. Bergman. The method of proper orthogonal decomposition for dynamical characterization and order reduction of mechanical systems: an overview,. *Nonlinear Dynamics*, 41:141–170, 2005.

[25] N. Kinkaid, O. O'Reilly, and P. Papadopoulos. Automotive disc brake squeal. *Journal of Sound and Vibration*, 267,1:105–166, 2003.

[26] P. Lancaster. *Lambda–matrices and Vibrating Systems*. Dover Publications, Mineola, N.Y., 2002.

[27] R. B. Lehoucq, K. Maschhoff, D. Sorensen, and C. Yang. *ARPACK Software Package*. Rice University, 1996.

[28] G. Liles. Analysis of disc brake squeal using finite element methods. Technical Paper 891150, SAE, 1989. doi:10.4271/891150.

[29] D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann. Structured polynomial eigenvalue problems: Good vibrations from good linearizations. *SIAM J. Matrix Anal. Appl.*, 28:1029–1051, 2006.

[30] The MathWorks Inc., Natick, Massachusetts. *MATLAB version R2014b*, 2014.

[31] V. Mehrmann and C. Schröder. Nonlinear eigenvalue and frequency response problems in industrial practice. *Journal of Mathematics in Industry*, 1(1):7, 2011.

[32] V. Mehrmann and H. Voss. Nonlinear eigenvalue problems: A challenge for modern eigenvalue methods. *GAMM-Mitteilungen*, 27:121–152, 2005.

[33] H. Ouyang, W. Nack, Y. Yuan, and F. Chen. Numerical analysis of automotive disc brake squeal: a review. *Int. J. of Vehicle Noise and Vibration*, 1, 3/4:207–231, 2005.

[34] E. Rabinovicz. *Friction and Wear of Materials*. John Wiley and Son, New York, 1965.

[35] J. Sethna. *Statistical Mechanics: Entropy, Order Parameters and Complexity*. Oxford Master Series in Physics 14, 2006.

[36] G. Spelsberg-Korspeter and P. Hagedorn. Complex eigenvalue analysis and brake squeal: Traps, shortcomings and their removal. *SAE Int. J. Passeng. Cars - Mech. Syst.*, 5,4:1211–1216, 2012.

[37] G. W. Stewart and J.-G. Sun. *Matrix Perturbation Theory*. Academic Press, New York, 1990.

[38] F. Tisseur and K. Meerbergen. The quadratic eigenvalue problem. *SIAM Review*, 43 (2):235–286, 2001.

[39] L. N. Trefethen and M. Embree. *Spectra and pseudospectra, The behavior of nonnormal matrices and operators*. Princeton University Press, Princeton, NJ, 2005.

[40] R. Varga. *Geršgorin and his circles*. Springer-Verlag, Berlin, 2004.

[41] K. Veselić. *Damped osciallations of linear systems - a mathematical introduction*. Springer-Verlag, Berlin, 2011.

[42] S. Volkwein. *Proper Orthogonal Decomposition: Theory and Reduced-Order Modelling*. Lecture Notes, University of Konstanz Department of Mathematics and Statistics, 2013.