Nonlinear eigenvalue and frequency response problems in industrial practice[§]

V. Mehrmann[†] C. Schröder[†]

February 21, 2011

Abstract

We discuss the numerical solution of large scale nonlinear eigenvalue problems and frequency response problems that arise in the analysis, simulation and optimization of acoustic fields. We report about the cooperation with the company SFE in Berlin. We present the challenges in the current industrial problems and the state-of-the-art of current methods. The difficulties that arise with current off-the-shelf methods are discussed and several industrial examples are presented. It is documented that industrial cooperation is by no means a one-way street of transfer from academia to industry but the challenges arising in industrial practice also lead to new mathematical questions which actually change the mathematical theory and methods.

Keywords: nonlinear eigenvalue problem, frequency response problem, complex symmetric linear system, block-Arnoldi method, acoustic field computation, automotive industry.

AMS classification: 65F18, 15A18.

1 Introduction

Traffic noise emissions by transport vehicles, such as cars, trains or airplanes are one of the key factors restricting the quality of life in urban areas. Acoustic waves in dynamically moving vehicles arise from many different sources, such as e.g. the noise of engines or the vibrations of the structure due to external excitations like road contact or head wind. The reduction of such noise emissions is therefore an important factor in the design and the economic success of new products.

To minimize potential noise emissions already in an early design phase of new products (avoiding the costly production of prototypes), it is necessary to use simulation, optimization and control techniques based on mathematical models of the vehicle. Performing these tasks requires mathematical models that describe the acoustic field associated with the structure of the complete vehicle including its interaction with the environment such as, e.g., the air in and around the vehicle or the road or rail contact. Furthermore, these models must allow to identify and influence potential noise sources, and as a vision for the future they must allow to minimize the emissions.

Although much research is carried out in universities and industrial research and development departments, the model based minimization of the noise emissions of a complete car or train (not to mention airplane) including the majority of external and internal excitations is still a vision for the future. To achieve this, a joint effort is needed that includes the identification of sources, the construction of adequate mathematical models that incorporate all possible sources for acoustic waves, the analysis of these models, concerning robustness of their descriptions and their potential for optimization techniques, the development of numerical methods for the simulation and optimization of these models, as well as the implementation of these techniques as production software on modern high performance computers for industrial use.

[†]Institut für Mathematik, Ma 4-5, TU Berlin, Straße des 17. Juni 136, D-10623 Berlin, FRG; email: {mehrmann,schroed}@math.tu-berlin.de

[§]Research supported by *Deutsche Forschungsgemeinschaft*, via the DFG Research Center MATHEON, Mathematics for Key Technologies, in Berlin and by SFE GmbH, Voltastr. 5, 13355 Berlin.

1.1 Modeling

The modeling of acoustic fields inside or outside of dynamically moving vehicles typically uses coupled systems of partial differential equations (PDEs), e.g. for the generation of noise by vibrating parts, surface contact, engine noise or head wind. These methods are well established in the engineering community using commercial finite element packages [4, 35]. However, the techniques for the resulting systems of PDEs still have many deficiencies, in particular the development of appropriate solvers for the solution of the linear and nonlinear systems and eigenvalue problems that have to be solved after discretization. These turn out to be extremely large and ill-conditioned when a reasonably fine 3D model is used; they may consist of hundreds of millions of equations.

An even greater challenge is to use these models and methods within an optimization loop. There is no chance to use classical off-the-shelf optimization methods for these problems, the problem size is just too large. Instead, one currently applies model reduction techniques [5, 10] to approximate the given fine model that is used for the simulation by a rather crude model that is used in the optimization.

To make such an approach viable within a design environment, where not only the geometry and topology of the vehicle and its material parameters are subject to changes, but also the interaction with the environment is rather complicated, it is necessary to develop integrated mathematical techniques for the modeling, simulation, and optimization that make as much as possible use of the properties of the underlying physical model and to transfer this into numerical methods and production codes.

1.2 Content of the paper

In this paper we will discuss some of the work and some of the challenges in a long-term cooperation with the company SFE GmbH in Berlin, Germany, which produces software for the simulation and optimization of acoustic fields.

The cooperation involves the development (and implementation on current high-performance computers) of numerical methods for the solution of linear systems and nonlinear eigenvalue problems arising from discretized partial differential equations modeling noise emissions of cars and trains.

The paper is organized as follows.

Section 2 briefly describes the modeling of acoustic fields inside a car as coupling of fluid and structure vibrations. In Section 3 we study the direct frequency response problem, i.e., the numerical solution of a series of large, sparse, complex symmetric linear systems with a varying parameter. We will address in particular in-core/out-of-core storage methods, preconditioning and parallel execution aspects.

Section 4 treats modal reduction to reduce the dimension of the generated models. This requires the computation of all eigenvalues of a large sparse nonlinear eigenvalue problems in a given region of the complex plane.

The described problems and results demonstrate many challenges in the transfer of state-of-the-art numerical methods into the industrial practice, and show that there is mutual benefit from such a cooperation between industry and academia.

2 Mathematical modeling of interior car acoustics

To model the propagation of acoustic waves inside a car, as illustrated in Figure 1, the 3-D lossless wave equation is used, which can be obtained from the continuity equation (conservation of mass)

$$\frac{\partial \tilde{\rho}}{\partial t} + \nabla(\tilde{\rho} v) = 0, \tag{1}$$

together with the Euler equations of fluid dynamics

$$\tilde{\rho}(\frac{\partial v}{\partial t} + (v \cdot \nabla)v) = -\nabla \tilde{p}. \tag{2}$$

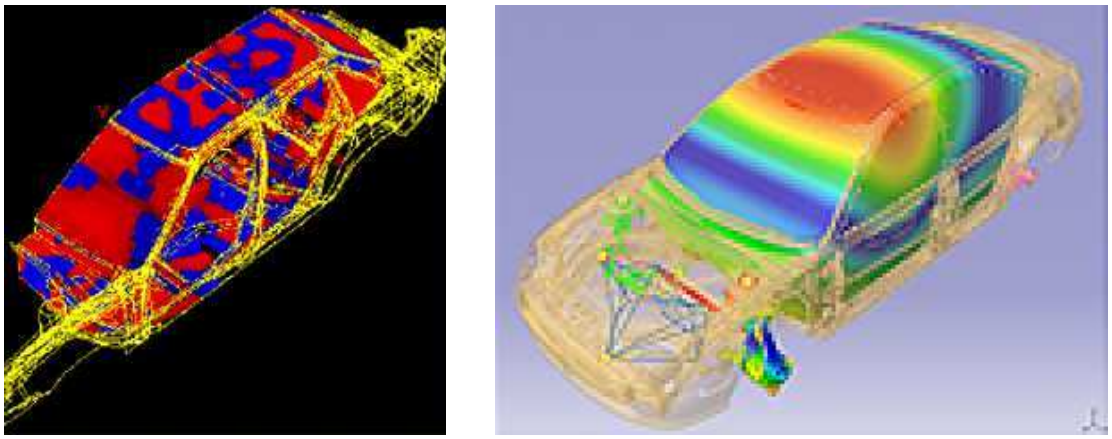

Figure 1: (left) FEM model of the car structure and (right) of the fluid

Here $v = v(x;y;z;t)$ is the particle velocity, $\tilde{\rho} = \tilde{\rho}(x;y;z;t)$ is the particle density, and $\tilde{p} = \tilde{p}(x;y;z;t)$ denotes the pressure, depending on cartesian coordinates x, y, z and time t. Under the simplifying assumptions that there is no temperature change, that the fluid is inviscid, i.e., no shear forces occur, that the influence of external forces is restricted to those coming from displacements of the structure at the boundaries, that the fluid is adiabatic, i.e., there is no heat exchange during compression, that we have an ideal gas, i.e., $\rho = \frac{p}{c^2}$, where c is the speed of sound, and finally that $(v \cdot \nabla)v$ and $\rho\frac{\partial v}{\partial t}$ are small, we can make the Taylor expansions

$$\begin{aligned} \tilde{p} &= p_0 + p(x;y;z;t), \text{ with } p_0 \gg p \ (p_0 = 10^6 p), \\ \tilde{\rho} &= \rho_0 + \rho(x;y;z;t), \text{ with } \rho_0 \gg \rho. \end{aligned}$$

Then from the Euler equation we obtain $\rho_0(\frac{\partial v}{\partial t}) = -\nabla p$ or by differentiation $\rho_0\nabla(\frac{\partial v}{\partial t}) = -\Delta p$, and from the continuity equation we get $\frac{\partial \rho}{\partial t} + \rho_0 \nabla v = 0$. Altogether we obtain the highly simplified system of partial differential equations

$$\frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} + \rho_0 \nabla \frac{\partial v}{\partial t} = \Delta p + \rho_0 \nabla \frac{\partial v}{\partial t} = 0$$

to which we have to add appropriate initial and boundary conditions. In a closed structure like the interior of the car, we can use the boundary conditions that are obtained from the displacement of the structure to obtain the fluid structure interaction, in an open structure like the acoustic waves emitted by the car to the outside, we have to incorporate appropriate far field conditions or transparent boundary conditions [41].

Let u be the vector of displacements of the structure on the surface. Then $v = \frac{\partial u}{\partial t}$ and thus with the outer normal ν we get

$$\nu\rho_0 \frac{\partial^2 u}{\partial t^2} = -\nu \nabla p.$$

To obtain a variational formulation, we multiply the equations with a test function w, and use integration by parts by integrating over control volumes V with surface elements S. This gives

$$\int_V \frac{1}{c^2} w \frac{\partial^2 p}{\partial t^2}\, dV + \int_V (\nabla w)\nabla p\, dV = -\rho_0 \int_S \nu w \frac{\partial^2 u}{\partial t^2}\, dS,$$

or equivalently

$$\int_V \frac{1}{\rho_0 c^2} w \frac{\partial^2 p}{\partial t^2}\, dV + \int_V \frac{1}{\rho_0} (\nabla w)\nabla p\, dV = -\int_S \nu w \frac{\partial^2 u}{\partial t^2}\, dS.$$

One of the most difficult tasks in the modeling is the incorporation of appropriate damping and absorption, because this depends in a rather complicated way on the materials used inside the

car, the surface geometry of the interior and this is also one of the terms, where the influence of the optimization is strong.

In the model treated here, damping and absorption was realized by an additional term

$$\int_S w \frac{r}{\rho_0^2 c^2} \frac{\partial p}{\partial t}\, dS,$$

where $r = r(\alpha)$ is depending on the damping properties of the material (described by a parameter vector α). We then obtain

$$\begin{aligned} \int_V \frac{1}{\rho_0 c^2} w \frac{\partial^2}{\partial t^2} p \, dV &+ \int_S w \frac{r}{\rho_0^2 c^2} \frac{\partial p}{\partial t}\, dS \\ &+ \int_V \frac{1}{\rho_0} (\nabla w) \nabla p dV = -\int_S \nu w \frac{\partial^2 u}{t^2}\, dS. \end{aligned}$$

To this variational formulation we can directly apply finite element discretization in space. This leads (in time) to a second order system of implicit differential equations in descriptor form

$$M_f \ddot{p}_d + D_f \dot{p}_d + K_f p_d + D_{sf} \ddot{u}_d = 0, \tag{3}$$

where $M_f = M_f^T$ is a positive definite mass matrix, $K_f = K_f^T$ is a positive definite stiffness matrix, $D_f(\alpha)$ is a symmetric positive semidefinite damping/absorption matrix, and D_{sf} describes the fluid structure coupling.

It should be noted that for the fluid model, the finite element discretization applied to the weak form of the partial differential equation is used. In contrast to this in the model describing the vibration of the structure, in industrial and engineering practice, typically a direct discrete finite element model is employed. This leads to one of the challenges that we will discuss below.

For the displacement vector u_d in the different finite elements of the discretized structure, assuming linear material laws, one directly obtains a linear second order system of differential-algebraic equations given by

$$M_s \ddot{u}_d + D_s \dot{u}_d + K_s u_d = f_e + f_p, \tag{4}$$

where f_e is a (discrete) external load and f_p is the pressure load. Here M_s is a symmetric positive definite mass matrix, and D_s is a symmetric positive semidefinite damping matrix; both are real. The matrix K_s has the form $K_s = K_1(\omega) + \imath K_2$ with real symmetric K_1, K_2, where K_1 is the positive semidefinite stiffness matrix. It is often frequency dependent to model nonlinear material behavior. The matrix K_2 models hysteretic damping, i.e., damping that is proportional to the displacement (instead of the velocity), but is in phase with velocity [13]. Typically, the matrix K_2 is of very small rank.

Although M_s is positive definite in theory, the matrix encountered in practice is highly singular due to the fact that rotational masses are omitted. On the positive side, it is block diagonal with small blocks.

It remains to further discuss the coupling of the fluid part and the structure part of the system. The term f_p originates from the pressure load $F_p = \int_S p\nu \, dS$. Finite element modeling/discretization yields $f_p = D_{sf}^T p_d$ and hence,

$$M_s \ddot{u}_d + D_s \dot{u}_d + K_s u_d - D_{sf}^T p_d = f_e. \tag{5}$$

If we combine all the equations then we obtain a second order system of differential-algebraic equations

$$\begin{bmatrix} M_s & 0 \\ D_{sf}^T & M_f \end{bmatrix} \begin{bmatrix} \ddot{u}_d \\ \ddot{p}_d \end{bmatrix} + \begin{bmatrix} D_s & 0 \\ 0 & D_f \end{bmatrix} \begin{bmatrix} \dot{u}_d \\ \dot{p}_d \end{bmatrix} + \begin{bmatrix} K_s(\omega) & -D_{sf}^T \\ 0 & K_f \end{bmatrix} \begin{bmatrix} u_d \\ p_d \end{bmatrix} = \begin{bmatrix} f_s \\ 0 \end{bmatrix}. \tag{6}$$

Typically in this model the format of M_s is a factor $1000 - 10\,000$ larger than the format of M_f, the mass and stiffness matrices are highly dependent on the geometry and topology of the car,

and the type of finite elements that are used. Since the structure is essentially modeled with a fine (pretty much uniform) mesh, the matrices have dimensions of several millions. The stiffness matrix depends also on the material parameters, while the damping and coupling matrices depend on geometry, topology and material parameters.

3 Direct Frequency Response

Usually one of the first tasks in the analysis of acoustic fields is to solve the frequency response problem, i.e., to simulate the behavior of the acoustic field under excitations (typically of the structure). The classical approach for the frequency response analysis of linear vibrational systems is to perform a Fourier ansatz

$$\begin{bmatrix} u_d \\ p_d \end{bmatrix} = \begin{bmatrix} \hat{u} \\ \hat{p} \end{bmatrix} e^{\imath\omega t}, \; f_s = \hat{f}(\omega) e^{\imath\omega t},$$

which then leads to a frequency dependent linear system of equations

$$\left(-\omega^2 \begin{bmatrix} M_s & 0 \\ D_{sf}^T & M_f \end{bmatrix} + \imath\omega \begin{bmatrix} D_s & 0 \\ 0 & D_f \end{bmatrix} + \begin{bmatrix} K_s(\omega) & -D_{sf}^T \\ 0 & K_f \end{bmatrix} \right) \begin{bmatrix} \hat{u}(\omega) \\ \hat{p}(\omega) \end{bmatrix} = \begin{bmatrix} \hat{f}(\omega) \\ 0 \end{bmatrix}. \tag{7}$$

This linear system, which (for a reasonable structure) has several millions of equations, typically has to be solved for a large frequency range $\omega = 0, \ldots, \; 1000 \; Hz$ in small frequency steps and, depending on the type of excitation, for many right hand sides (load vectors).

Based on the frequency response analysis it is then possible to detect places where the excitation leads to large noise emissions (hot spots), and this approach can be used to improve or even optimize the frequency response behavior within an optimization loop.

Before we discuss solution methods let us introduce an important modification of (7). For nonzero frequencies we may multiply the second block row by ω^{-1} and introduce the new variable $\hat{v}(\omega) = \omega \hat{p}(\omega)$. Then we obtain a new linear system

$$\left(-\omega^2 \begin{bmatrix} M_s & 0 \\ 0 & M_f \end{bmatrix} + \imath\omega \begin{bmatrix} D_s & \imath D_{sf}^T \\ \imath D_{sf}^T & D_f \end{bmatrix} + \begin{bmatrix} K_s(\omega) & 0 \\ 0 & K_f \end{bmatrix} \right) \begin{bmatrix} \hat{u}(\omega) \\ \hat{v}(\omega) \end{bmatrix} = \begin{bmatrix} \hat{f}(\omega) \\ 0 \end{bmatrix}, \tag{8}$$

which now has complex symmetric coefficients and, to simplify notation, we write this as

$$P(\omega)x(\omega) := \left(-\omega^2 M + \imath\omega D + K(\omega) \right) x(\omega) = b(\omega). \tag{9}$$

Since this system of equations has to be solved for many right hand sides and over a large frequency range, and all this within an optimization loop, a very efficient solver is necessary. This solver has to be able to recycle information from nearby problems (when stepping through the frequency range and modifying parameters in an optimization) and has to work efficiently in a modern multi-processor, multi-core hardware environment. Altogether, this is really a lot to ask for from a linear system solver and makes the use of black box solvers extremely difficult.

3.1 State-of-the-art in linear system solvers

Modern techniques for the solution of linear systems in industrial applications are typically a combination of methods, that use the best of each of the various available classes of methods.

The first class of methods are highly efficient direct solution methods that use Gaussian elimination with partial pivoting combined with other techniques from graph theory to achieve optimal performance, make efficient use of the sparsity structure to avoid fill-in, and save storage. Many of them are even implemented for the use on multiprocessor machines. Well known packages include UMFPACK [17], PARDISO [40], MUMPS [3], WSMP [26], or the HSL collection (e.g., HSL_MA78 [36]) to name a few. In view of possibly many right hand sides they are a clear option, despite the fact that for the size and type of problems considered in industrial practice, the

storage demands are so high that they typically can only work in an out-of-core setting, i.e., the matrix factors are stored on hard disk instead of main memory. Another difficulty is that a new factorization has to be computed for every frequency or every modification in the optimization loop. Finally also the bad scaling and the fact that the problems get increasingly ill-conditioned for large frequencies presents a real challenge, because the desired solution accuracy may not be realizable. Some of the packages provide specialized routines for complex symmetric systems, which have the potential to half the storage and computational requirements, but may also increase the complexity of pivoting strategies.

The second class of methods are iterative methods of the Krylov subspace type like the generalized minimal residual method (GMRES) [39], the Bi-conjugate gradient method in its stabilized form (BICGSTAB) [50], or the quasi-minimal residual method (QMR) [23]. A variant of the conjugate gradient method for complex symmetric systems like (9) is CSYM [14]. In principle these methods would be very good candidates, since the linear systems are sparse and matrix vector multiplications are relatively cheap. Also, if $\hat{f}$ and K are independent of ω, then one can exploit the shift invariance property of Krylov subspaces [25, 45]. However, often $\hat{f}$ or K do depend on ω and without a good preconditioner the convergence of iterative methods (in particular for large frequencies) is dramatically slow or not realizable.

The third class of potential methods (exploiting the fact that there is a partial differential equation in the background) are multigrid or multilevel methods [27, 31, 32]. Unfortunately, despite the fact that they are efficiently used in the solution of wave propagation problems, such as Helmholtz or Maxwell equations [19], they cannot be used in a simple way in the described acoustic field problems, since the discrete FEM modeling of the structure does not provide a nice hierarchy of basis functions. If such a hierarchy was available then these methods would lead to very good preconditioners for the Krylov subspace methods. However, with the current state of discrete FEM modeling (being engineering practice for decades) it is extremely difficult to incorporate them into the current code solvers. This is even true for the algebraic multigrid methods like the Ruge-Stüben method [37] or hypre [20], due to the difficulty of having to solve for a large frequency range, many right hand sides and all this inside an optimization loop. It is needless to say that also the incorporation of these solvers in a multiprocessor, multi-core industrial software tool is another challenge.

The fourth class of methods which have made tremendous progress in the last decades are the adaptive finite element methods, see e.g. [9]. These refine the computational grid according to a priori and a posteriori error estimates and if implemented properly avoid globally fine meshes and therefore the extremely large linear systems and eigenvalue problems in the first place. However, the construction, analysis, and implementation of such methods for their use inside an industrial package, including the treatment of a full frequency range and many right hand sides, is still in its infancy and requires a major research effort which fortunately is currently addressed in several research projects world-wide.

When our cooperation with SFE started, we essentially made the above analysis of the available methods and realized the following major obstacles.

- The problems are badly scaled and get increasingly ill-conditioned when ω grows;
- for some parameter constellations the system becomes exactly singular with possibly inconsistent right hand sides;
- typically there are many right-hand sides;
- classical off-the-shelf iterative methods do not work well, or fail completely;
- direct solution methods have to work by storing the factors in out-of-core memory and cannot easily recycle information from previous frequency or optimization steps;
- no multilevel or adaptive grid refinement is applicable, the methods must be matrix based;
- the matrices M, K_1 are often slightly indefinite because of rounding errors during their creation.

3.2 What did we do?

With all the obstacles of the problem and all the deficiencies of the current methods, the development and implementation of an appropriate linear solver for acoustic field frequency response problems is a major challenge. Furthermore, as in all industrial projects, there was a deadline. In such a circumstance, the only possibility is to compromise between the optimality of a given method for a given problem, the provability of success, and the practical needs in the industrial environment.

Our compromise, since we had to deal with given matrices (as opposed to partial differential equations), was to develop and implement (in the SFE software environment) a preconditioned subspace recycling Krylov subspace method which has the following features:

The initial search space for any frequency ω is chosen to contain the solutions for the last few frequencies, an approach also used in [21]. For the preconditioner, whose construction and application represent the dominant cost of the algorithm, we had to fulfill the requirement that it had to run in a distributive setting, store its results out-of-core, support complex arithmetic, in particular support complex symmetric systems. This is a tall order to ask for but fortunately we were able to base the preconditioner on existing software, using the complex LDL^T factorization code of the MUMPS package [3]. The preconditioner was constructed from the symmetric real part of the linear system (9), i.e., from

$$\tilde{P}(\omega) := -\omega^2 \begin{bmatrix} M_s & 0 \\ 0 & M_f \end{bmatrix} + \begin{bmatrix} K_1 & 0 \\ 0 & K_f \end{bmatrix}.$$

For small values of ω up to ca $60 \cdot 2\pi$ corresponding to a frequency of 60 Hz, typically only $3-6$ iteration steps per frequency were necessary and the preconditioner could be kept unchanged for the whole frequency range. But the number of iteration steps per frequency grows substantially the larger ω gets, so that more and more new preconditioners have to be computed. For large frequencies close to 1000 Hz, the preconditioned Krylov method tends to be extremely slow or not convergent at all. As a consequence we constructed a hybrid method, where from a certain frequency on MUMPS is used as a direct solver for the system with the full matrix $P(\omega)$. In Figure 2 we present a comparison of our solver with the solver from NX Nastran for a model with 219432 degrees of freedom, solved on 2 $\times$ AMD Opteron DualCore 2600MHz with 8GB RAM under Suse Linux 10.0. Furthermore, it was decided to make use of the multi-core environment and to treat several linear systems simultaneously in a distributed fashion. One group of processors solves the systems corresponding to the lower frequencies using the described iterative recycling method. Meanwhile, the other processor groups treat the high frequency systems with a full direct solve for each value of ω.

3.3 Evaluation of our approach

Since there are commercial packages available, a question that may arise is why we need new numerical methods and solvers in the first place. From the point of view of the company SFE this is clear, they wanted their own solver in their product, and the solution method including the implementation was satisfactory for their needs. But in many respects the current methods are unsatisfactory from a scientific point of view. The discussed industrial problems present some of the current grant challenges in the field and the commercial packages are by no means able to solve these problems in a completely satisfactory way. To always get a more-or-less reasonable solution, and to be efficient, they often compromise for accuracy of the results which is certainly inadequate from a scientific point of view. Thus the cooperation with the industrial partner triggered new research questions for the academic world.

It became clear during the project that the concept of *recycling* in iterative methods, i.e., the reuse of information that was already computed for other frequencies or in the course of an optimization procedure is not well-enough understood. As a consequence, motivated by the project with SFE, a new research project [2] in the DFG Research Center MATHEON was started to further

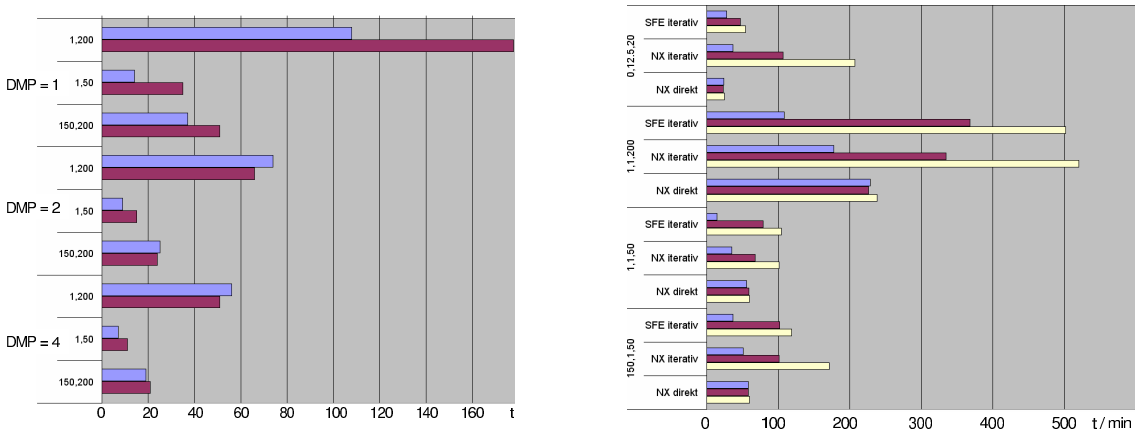

Figure 2: (left) Light blue bars: our solver, dark red bars: NX Nastran; on 1, 2, or 4 Cores (DMP); for the range of frequencies $(1Hz, \ldots, 200Hz)$, $(1Hz, \ldots, 50Hz)$, $(150Hz, \ldots, 200Hz)$ in $1Hz$ steps; with 1 right hand side (RHS). (right) Light blue bars: 1 RHS, dark red bars: 5 RHS, light yellow bars: 10 RHS; in every 3 consecutive bar groups: top group: our solver, middle group: Nastran iterative solver, bottom group: Nastran direct solver; top 3 bar groups: Frequency range: $(1Hz, \ldots, 200Hz)$ in $12.5Hz$ steps, 2nd 3 bar groups: $(1Hz, \ldots, 200Hz)$ in $1Hz$ steps, 3rd 3 bar groups: $(1Hz, \ldots, 50Hz)$ in $1Hz$ steps, bottom 3 bar groups: $(150Hz, \ldots, 200Hz)$ in $1Hz$ steps; all using 1 Core.

investigate the basic mathematical principles and to understand how they can be incorporated into new efficient methods [24].

As a second major obstacle, we identified the use of discrete FEM modeling in structural engineering. It would be much easier if adaptive FEM would be usable for the discussed class of problems, and it would also be good to have a grid hierarchy that allows the use of efficient multilevel preconditioners [27, 31, 32]. Despite the high research activity in this field this is a major challenge for the described problem classes and almost no analysis or methods are available. We will discuss this in more detail in the section devoted to eigenvalue computation, but again here we see a need for a stronger cooperation between academia and industry in this area of structural engineering, to transfer new ideas that are developed now into the industrial practice.

4 Modal reduction

We have seen in the previous section that due to the fine mesh used for the discrete FEM modeling of the structure, the linear systems have very large dimensions. Typically, however, one is interested in damping the low frequencies associated with the eigenvalues in the neighborhood of 0 (and the imaginary axis) of the complex symmetric matrix function

$$Q(\lambda) := \lambda^2 M + \lambda D + K = Q(\lambda)^T, \tag{10}$$

with M, D, K as in (9). The methods discussed below assume that P is quadratic in λ for the eigenvalue computation, i.e. the nonlinear dependency of K on the frequency is ignored. One easily verifies that if λ_0 is an eigenvalue of $Q(\lambda)$, i.e., $P(\lambda_0)x = 0$, then x^T is the corresponding left eigenvector but essentially no other general properties of the eigenvalue problem are available, in particular, there is unfortunately no immediate variational property for the eigenvalues/eigenfunctions available, as there is for the undamped case [33].

Since one is interested only in part of the spectrum, a natural idea is to identify a space associated with eigenvectors and generalized eigenvectors associated with the important eigenvalues, and to project the problem into this subspace. This is a *model reduction* approach called *modal re-*

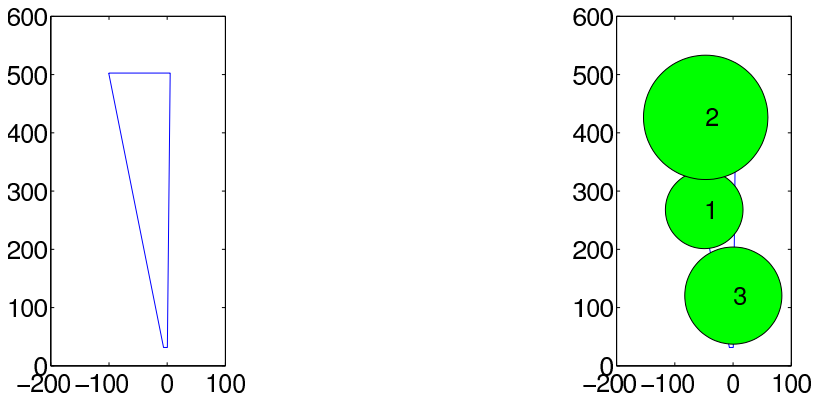

Figure 3: (left) A typical trapezoidal region within which all eigenvalues are sought. (right) The trapezoidal region covered with a few circles.

duction which, if efficiently implemented, can save a huge amount of computing time and storage, despite the fact that it is partially heuristic.

4.1 Complex symmetric quadratic eigenvalue problems

In view of the desire to do modal reduction, another task in the cooperation with SFE GmbH was the computation of a small number, say ℓ, eigenvalues in a trapezoidal region near zero, see Figure 3 (a), and the corresponding subspace $\mathcal{S}_\ell$ spanned by the eigenvectors and generalized eigenvectors associated with these eigenvalues, i.e., $\mathcal{S}_\ell = \text{span}\,\{s_1, \ldots, s_\ell\}$, where the s_i form an orthonormal basis of this subspace, i.e., $S_\ell = [s_1, \ldots, s_\ell]$ is an isometry.

The projected system would have the form

$$Q_\ell(\lambda) := \lambda^2 M_\ell + \lambda D_\ell + K_\ell := \lambda^2 S_\ell^T M S_\ell + \lambda S_\ell^T D S_\ell + S_\ell^T K S_\ell \tag{11}$$

and would still be complex symmetric. In the context of model reduction, the requirements for the reduced model are that the projected system is a good approximation to the large scale system for a large frequency range, and also for a large set of parameter variations. Furthermore, it has to be of small enough dimension, so that classical methods for nonlinear nonsmooth optimization can be applied. Again this is a lot to ask, since currently for large scale problems there are really no methods available that guarantee to obtain all the eigenvalues of (10) in a specified region of the complex plane. For small dense problems one could employ the sign function method, [8, 28] but this would require storing full dense inverses of matrices of the given class, which is certainly not possible in the described acoustic field problems.

The currently used industrial techniques typically solve a simplified problem, e.g., the eigenvalues and associated eigenvectors in the desired region for the undamped/uncoupled problem are used for the projection or an algebraic multilevel substructuring method (AMLS) is used that exploits the structure of the matrices as they arise form the discrete FEM [11]. All these techniques are partially heuristic, since there is no guarantee that all the desired eigenvalues are captured or that the eigenvalues from the projected system are close to those of the original physical system or of the FEM discretized system, i.e., in general, currently no error bounds are available. To generate such error bounds is a major challenge for the academic community and research in this direction has started in the project [1] with first results in [15, 34].

Furthermore, the numerical solution of the large scale nonlinear eigenvalue problem (10) itself also presents a mathematical challenge, for several reasons.

First of all the mass matrix is singular, so the problem has eigenvalues at ∞ which often cause convergence problems for the iterative methods, second of all, there are also several (typically six)

eigenvalues at 0, corresponding to the six free degrees of motion, three each for translation and rotation, of the whole structure. In industrial practice the eigensolver is furthermore also often used for model verification, by checking whether there are exactly these 6 eigenvalues at 0. Then, if the model is flawed, the singularity may be even higher.

The major challenge, however, is that we want the eigenvalues near 0 and it is well known that classical iterative methods for large sparse eigenvalue problems, like the implicitly restarted Arnoldi method [29], typically converge fast only to the eigenvalues at the periphery of the spectrum [38]. Thus, either a shift-and-invert technique that transforms the problem and maps the desired eigenvalue to the periphery is necessary or other techniques like Newton's method [43, 44] or the Jacobi-Davidson method [46, 47] have to be used. All these methods require the solution of linear systems of the form $(\hat{\lambda}^2 M + \hat{\lambda} D + K)x = b$ with a given shift-point $\hat{\lambda}$. But this is the problem that we wanted to avoid in the first place and hence a vicious circle is closed.

4.2 What did we do?

Having assessed the various options and their potential advantages and drawbacks, we decided to develop a new method based on the following concepts and to implement it into the SFE environment.

First of all, since the methods that work directly on the quadratic eigenvalue problems, e.g., [12, 33, 49], are not yet as mature as those for linear eigenvalue problems, the problem is linearized by introducing a new variable λx and turning (10) from a quadratic into a linear eigenvalue problem.

Since there is no method available that is guaranteed to find all eigenvalues in a region at feasible costs, we had to revert to a partially heuristic approach. We are looking for eigenvalues in the vicinity of target points which are scattered inside the region of interest. To find eigenvalues close to a particular target σ, also called a shift, we use the shift-linearize-invert-ansatz [49] that requires computation of the largest eigenvalues of the matrix

$$\mathcal{A} := \begin{bmatrix} Q(\sigma)^{-1} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} 2\sigma M + D & M \\ -I & 0 \end{bmatrix}.$$

This is done by the block Arnoldi method [6] which requires the application of the matrix $\mathcal{A}$ to given blocks of vectors during the generation of the Krylov subspace. This involves a solution of a linear system with the complex symmetric matrix $Q(\sigma)$. We again use the direct solver MUMPS to compute a complex LDL^T factorization.

The block Arnoldi method searches for eigenvectors of $\mathcal{A}$ within the block Krylov subspace

$$\mathcal{K}_m(\mathcal{A}, B) := span\{B, \mathcal{A}B, \mathcal{A}^2 B, \ldots, \mathcal{A}^{m-1} B\}, \tag{12}$$

which is known to contain, for increasing m, increasingly good approximations to eigenvectors corresponding to eigenvalues of $\mathcal{A}$ of maximum absolute value — which correspond to the eigenvalues of $Q(\lambda)$ closest to σ. In the classical Arnoldi method, the starting block B is a vector. Otherwise, when B consists of several, say n_b, columns, the resulting method is the block Arnoldi method. The matrix B can be chosen in an (almost) arbitrary way, but if eigenvector or subspace approximations are known, the use of these will speed up convergence. Among the advantages of using the block method are that clusters of eigenvalues are handled better [30]. Furthermore, all vector-vector operations become matrix-matrix operations which can be implemented much more efficiently by use of BLAS level 3 routines [18] and the matrix-vector multiplication with the large matrix $\mathcal{A}$ can be carried out with a block of vectors. Furthermore, the heuristic part of the algorithm is more reliable in the block case, see below.

In the actual implementation of the block Arnoldi algorithm the terms $\mathcal{A}^i B$ are never explicitly formed. Instead an orthonormal basis $Q_m = [V_1, V_2, \ldots, V_m]$ of $\mathcal{K}_m(\mathcal{A}, B)$ is generated by setting V_1 to be an orthonormal basis of $span(B)$ followed by iteratively forming V_{i+1} by orthonormalization of $\mathcal{A}V_i$ against $V_1, \ldots, V_i$. Classically, a variant of the Gram-Schmidt method is used for this task [29], but any other orthonormalization procedure may be employed. We use block Householder reflections [7, 42] that are very stable and attain high efficiency by using BLAS-3

operations. Collecting all the V_i and the orthonormalization coefficients $H_{i,j}$ results in the well known Arnoldi relation $\mathcal{A}Q_m = Q_m H_m + V_{i+1} H_{m+1,m} E_m^T$ with an $(m \times m)$–block Hessenberg matrix $H_m = (H_{i,j})_{i,j=1}^m$. The eigenvalues of H_m, called Ritz values, are used as approximations of eigenvalues of $\mathcal{A}$. Likewise, the eigenvectors of H_m, multiplied with Q_m, are used as approximate eigenvectors of $\mathcal{A}$.

A drawback of the block Arnoldi method is the necessity to store the basis Q_m which grows for increasing m. A typical way out of this problem is to restart the algorithm at some point. Instead of restarting with a single new starting block it is possible to restart with a whole Arnoldi relation. For the vector Arnoldi method, two such implicit restarting schemes are common, one using a filter polynomial [29] and one using a reordering of the Schur form of H_m [48]. We are using the latter approach as the first one does not elegantly generalize to the block case [7].

The heuristic part of our method is that we assume that the block Arnoldi method really finds all eigenvalues in a circle around the shift σ. While quite often the eigenvalues do indeed appear and converge in the order of the distance to the shift, it is not rare that one or a group of eigenvalues converge slower than other farer away eigenvalues. However, in these cases usually the missed eigenvalues are present as (yet) unconverged Ritz values. Therefore, we use the unconverged Ritz value that is closest to the shift as the radius of a circle that is trusted to contain no missed eigenvalues.

In the implementation, we repeatedly run the block Arnoldi method for different shifts, possibly several at once in a distributed setting. Figure 3 (right) shows the situation after a few iterations. Large parts of the trapezoidal region are covered, leaving only some small remaining regions to be searched. New shifts are placed inside the largest such white regions, until the whole trapezoidal region of interest is covered by trusted circles.

Of course, with such a covering approach an eigenpair could be computed by more than one Arnoldi run for different shifts. For that reason the freshly discovered eigenpairs have to be checked for being copies of already previously found pairs. To achieve this we consider a new eigenpair (λ_*, x_*) being a copy, if $[x_*^T, \lambda_* x_*^T]^T$ is almost linearly dependent to the span of the vectors $[x_{\text{old}}^T, \lambda_{\text{old}} x_{\text{old}}^T]^T$ corresponding to every known eigenvalue λ_{old} sufficiently close to λ_*.

We applied the solver to a car model, discretized by a regular mesh of 35mm leading to 219432 degrees of freedom. We were looking for eigenvalues within the triangle bordered by the lines $\text{Im}(\lambda) > 20\text{Re}(\lambda)$, $\text{Im}(\lambda) > -20\text{Re}(\lambda)$, and $\text{Im}(\lambda) < f \cdot 2\pi$, for $f \in \{50, 100, 150, 250\}$. The following table lists the number of eigenvalues within these triangles, the number of used shifts, i.e., used matrix factorizations, and the number of overall Arnoldi iterations, i.e., Matrix-block products.

0Hz to	50Hz	100Hz	150Hz	200Hz	250Hz
number of found eigenvalues	20	52	129	217	346
number of shifts	1	3	3	5	7
number of iterations	39	139	136	294	437

The test was performed on a PC with an Intel Core2 Duo E6850 CPU clocked at 3.0GHz, with 4Gb RAM. One shift was addressed at a time using one processor. The blocksize was 5.

4.3 Evaluation of our approach

Our approach certainly does not use many novel ideas, but instead builds on mature and proven concepts. The method can run in a distributed setting processing several shifts at once, each one running in several processes. The matrix factorizations can be kept out-of-core. Moreover, since the basis vectors are explicitly orthogonalized by a stable Householder scheme, the generated block Arnoldi basis is orthonormal to working precision. This is in sharp contrast to the basis generated by the nonsymmetric Lanczos method which typically looses linear independence after enough iterations and leads to so-called spurious eigenvalues [16, 22].

In our experiments the heuristic choice of the radii of trusted circles worked very well. In thousands of test cases with problems from SFE as well as randomly generated problems, it happened only once that this approach missed an eigenvalue. This happened with a block size of

one, i.e. with the vector Arnoldi method. The method worked fine for this case if a block size of at lest two was used. In our implementation the default block size is 8.

Unfortunately, the improved robustness of the block-Arnoldi method compared with the standard Arnoldi method comes at the price of increased memory requirements to store the basis. This downside is somewhat mitigated, however, by the use of restarts.

As another downside of our approach, so far only quadratic eigenvalue problems can be solved, for truly nonlinear problems with ω dependence in K the methods is not directly applicable. On the other hand, the method can be applied to fluid or structure subsystems separately, or to the complete coupled system, and it tolerates a singular mass matrix.

5 Conclusions. The two-way street of industrial cooperation

We discussed the modeling, simulation and (at least the vision) of the optimization of acoustic fields. As an example of our industrial cooperation we studied frequency domain computations and modal reduction methods for the acoustic field inside a car, as well as the challenges of the resulting linear systems and eigenvalue problems. One lesson from the project was that the methods to be used in industrial practice cannot be constructed and implemented using textbook approaches. For instance, very often tricks have to be employed that increase the efficiency of the computation, but that lack a full mathematical understanding. This can lead to misunderstandings between engineers, programmers and mathematicians. A stronger communication and cooperation between these groups is necessary to address the described challenges. If this is achieved then all sides benefit from a cooperation. The work with SFE GmbH on the frequency response problem for interior acoustic field computation started out as a clear transfer project (a one way street), with the idea to transport the knowledge and know-how about current linear system and eigenvalue solver technologies available in the academic environment into an industrial software environment.

But as we have discussed, already early on in the cooperation a lot of new research topics appeared that could not be treated within the current project (two-way-street). Examples include recycling methods for a sequence of slowly changing linear systems, updating of preconditioners, guaranteed location of all eigenvalues inside a region of the complex plane, to name a few.

Acknowledgments

We thank H. Zimmer and A. Hilliges from our industrial partner SFE for the permission to use the figures obtained in the cooperation, and we thank Tobias Brüll, Elena Teidelt, Matthias Miltenberger for their help in the implementation.

References

[1] Adaptive solution of parametric eigenvalue problems for partial differential equations. Project C22 of DFG research center MATHEON in Berlin, url: `http://www.matheon.de/research/show_project.asp?id=126`.

[2] Numerical methods for large-scale parameter-dependent systems. Project C29 of DFG research center MATHEON in Berlin, url: `http://www.matheon.de/research/show_project.asp?id=172`.

[3] MUMPS: MUltifrontal Massively Parallel Solver. Users guide, November 2009. `http://graal.ens-lyon.fr/MUMPS`.

[4] ANSYS, Inc. *Programmer's manual for ANSYS*, 2007. http://www1.ansys.com/customer/content/documentation/110/ansys/aprog110.pdf.

[5] A. C. Antoulas. *Approximation of large-scale dynamical systems*, volume 6 of *Advances in Design and Control*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2005.

[6] W. E. Arnoldi. The principle of minimized iteration in the solution of the matrix eigenvalue problem. *Quart. Appl. Math.*, 9:17–29, 1951.

[7] J. Baglama. Augmented block Householder Arnoldi method. *Linear Algebra Appl.*, 429(10):2315–2334, 2008.

[8] Z. Bai and J. Demmel. Using the matrix sign function to compute invariant subspaces. *SIAM J. Matrix Anal. Appl.*, 19(1):205–225 (electronic), 1998.

[9] W. Bangerth and R. Rannacher. *Adaptive Finite Element Methods for Differential Equations*. ETH lecture series. Birkhäuser, Basel, Ch., 2003.

[10] P. Benner, V. Mehrmann, and D. Sorensen, editors. *Dimension Reduction of Large-Scale Systems*, volume 45 of *LNSCE*. Heidelberg, 2005. ISBN 3-540-24545-6.

[11] J. K. Bennighof. Adaptive multi-level substructuring method for acoustic radiation and scattering from complex structures. In A. J. Kalinowski, editor, *Computational Methods for Fluid/Structure Interaction*, volume 178, pages 25–38. AMSE, New York, Nov 1993.

[12] W.-J. Beyn. An integral method for solving nonlinear eigenvalue problems. *ArXiv e-prints*, March 2010.

[13] R.E.D. Bishop. The treatment of damping forces in vibration theory. *Jl R. aeronaut. Soc.*, 59:738, 1955.

[14] A. Bunse-Gerstner and R. Stöver. On a conjugate gradient-type method for solving complex symmetric linear systems. *Linear Algebra Appl.*, 287(1-3):105–123, 1999.

[15] C. Carstensen, J. Gedicke, V. Mehrmann, and A. Miedlar. An adaptive homotopy approach for non-selfadjoint eigenvalue problems. Preprint 718, DFG Research Center MATHEON, *Mathematics for key technologies* in Berlin, 2010.

[16] J. Cullum and R. A. Willoughby. Lanczos and the computation in specified intervals of the spectrum of large, sparse real symmetric matrices. In *Sparse Matrix Proceedings 1978 (Sympos. Sparse Matrix Comput., Knoxville, Tenn., 1978)*, pages 220–255. SIAM, Philadelphia, Pa., 1979.

[17] T. A. Davis. Algorithm 832: UMFPACK V4.3—an unsymmetric-pattern multifrontal method. *ACM Trans. Math. Software*, 30(2):196–199, 2004.

[18] J. J. Dongarra, J. Du Croz, S. Hammarling, and I. Duff. A set of level 3 Basic Linear Algebra Subprograms. *ACM Transactions on Mathematical Software*, 16(1):1–17, March 1990.

[19] O. G. Ernst. *Fast Numerical Solution of Exterior Helmholtz Problems with Radiation Boundary Condition by Imbedding*. PhD thesis, Stanford University, 1994.

[20] R. D. Falgout, J. E. Jones, and U. Meier Yang. The design and implementation of hypre, a library of parallel high performance preconditioners. In *Numerical solution of partial differential equations on parallel computers*, volume 51 of *Lect. Notes Comput. Sci. Eng.*, pages 267–294. Springer, Berlin, 2006.

[21] P. F. Fischer. Projection techniques for iterative solution of $A\underline{x} = \underline{b}$ with successive right-hand sides. *Comput. Methods Appl. Mech. Engrg.*, 163(1-4):193–204, 1998.

[22] R. W. Freund, M. H. Gutknecht, and N. M. Nachtigal. An implementation of the look-ahead Lanczos algorithm for non-Hermitian matrices. *SIAM J. Sci. Comput.*, 14(1):137–158, 1993.

[23] R. W. Freund and N. M. Nachtigal. QMR: a quasi-minimal residual method for non-Hermitian linear systems. *Numer. Math.*, 60(3):315–339, 1991.

[24] A. Gaul, M. H. Gutknecht, J. Liesen, and R. Nabben. Deflated and augmented krylov subspace methods: Basic facts and a breakdown-free deflated minres. Preprint 749, DFG Research Center MATHEON, Mathematics for key technologies in Berlin, 2011.

[25] G.-D. Gu and V. Simoncini. Numerical solution of parameter-dependent linear systems. *Numer. Linear Algebra Appl.*, 12(9):923–940, 2005.

[26] A. Gupta. WSMP: Watson Sparse Matrix Package, Part II. direct solution of general sparse systems. Ibm research report, IBM T. J. Watson Research Center, November 2000. http://www.alphaworks.ibm.com/tech/wsmp.

[27] W. Hackbusch. *Multigrid methods and applications*, volume 4 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 1985.

[28] N. J. Higham. *Functions of matrices. Theory and computation*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008.

[29] R. B. Lehoucq, D. C. Sorensen, and C. Yang. *ARPACK users' guide. Solution of large-scale eigenvalue problems with implicitly restarted Arnoldi methods*, volume 6 of *Software, Environments, and Tools*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1998.

[30] R.B. Lehoucq and K.J. Maschhoff. Implementation of an implicitly restarted block Arnoldi method. Preprint MCS-P6490297, Argonne National Laboratory, Argonne, IL, 1997.

[31] S. F. McCormick. *Multilevel adaptive methods for partial differential equations*, volume 6 of *Frontiers in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1989.

[32] S. F. McCormick. *Multilevel projection methods for partial differential equations*, volume 62 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.

[33] V. Mehrmann and H. Voss. Nonlinear eigenvalue problems: A challenge for modern eigenvalue methods. *Mitt. der Ges. f. Angewandte Mathematik und Mechanik*, 27:121–151, 2005.

[34] A. Miedlar. *Inexact Adaptive Finite Element Methods for Elliptic PDE Eigenvalue Problems*. Phd thesis, Technische Universität Berlin, Inst. f. Mathematik, 2011.

[35] MSC.Software Corporation. *MSC.Nastran 2005 Installation and Operations Guide*, 2004. http://www.mscsoftware.com/Products/CAE-Tools/MSC-Nastran.aspx.

[36] J. K. Reid and J. A. Scott. An efficient out-of-core multifrontal solver for large-scale unsymmetric element problems. Technical Report RAL-TR-2007-014, SFTC Rutherford Appleton Laboratory, 2007.

[37] J. W. Ruge and K. Stüben. Algebraic multigrid. In *Multigrid methods*, volume 3 of *Frontiers Appl. Math.*, pages 73–130. SIAM, Philadelphia, PA, 1987.

[38] Y. Saad. *Numerical methods for large eigenvalue problems*. Algorithms and Architectures for Advanced Scientific Computing. Manchester University Press, Manchester, 1992.

[39] Y. Saad and M. H. Schultz. GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Statist. Comput.*, 7(3):856–869, 1986.

[40] O. Schenk and K. Gärtner. Solving unsymmetric sparse systems of linear equations with PARDISO. *Future Generation Computer Systems*, 20(3):475 – 487, 2004.

[41] F. Schmidt and P. Deuflhard. Discrete Transparent Boundary Conditions for the Numerical Solution of Fresnel's Equation. *Computers Math. Applic.*, (29):53–76, 1995.

[42] R. Schreiber and C. Van Loan. A storage-efficient WY representation for products of Householder transformations. *SIAM J. Sci. Statist. Comput.*, 10(1):53–57, 1989.

[43] H. Schwetlick and K. Schreiber. A primal-dual Jacobi-Davidson-like method for nonlinear eigenvalue problems. Internal Report ZIH-IR-0613, Techn. Univ. Dresden, Zentrum für Informationsdienste und Hochleistungsrechnen, 2006.

[44] H. Schwetlick and K. Schreiber. Nonlinear Rayleigh functionals. *Linear Algebra and its Applications*, In Press, Corrected Proof, 2010. doi: 10.1016/j.laa.2010.06.048.

[45] V. Simoncini and F. Perotti. On the numerical solution of $(\lambda^2 A+\lambda B+C)x = b$ and application to structural dynamics. *SIAM J. Sci. Comput.*, 23(6):1875–1897 (electronic), 2002.

[46] G. L. G. Sleijpen, A. G. L. Booten, D. R. Fokkema, and H. A. Van der Vorst. Jacobi-Davidson type methods for generalized eigenproblems and polynomial eigenproblems. *BIT*, 36(3):595–633, 1996.

[47] G. L. G. Sleijpen and H. A. Van der Vorst. A Jacobi-Davidson iteration method for linear eigenvalue problems. *SIAM J. Matrix Anal. Appl.*, 17(2):401–425, 1996.

[48] G. W. Stewart. A Krylov-Schur algorithm for large eigenproblems. *SIAM J. Matrix Anal. Appl.*, 23(3):601–614, 2001/02.

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

[50] H. A. van der Vorst. Bi-CGSTAB: a fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems. *SIAM J. Sci. Statist. Comput.*, 13(2):631–644, 1992.