CISM International Centre for Mechanical Sciences 579
Courses and Lectures

Manfred Kaltenbacher *Editor*

Computational Acoustics

International Centre
for Mechanical Sciences

Springer

CISM International Centre for Mechanical Sciences

Courses and Lectures

Volume 579

The series presents lecture notes, monographs, edited works and proceedings in the field of Mechanics, Engineering, Computer Science and Applied Mathematics. Purpose of the series is to make known in the international scientific and technical community results obtained in some of the activities organized by CISM, the International Centre for Mechanical Sciences.

More information about this series at http://www.springer.com/series/76

Manfred Kaltenbacher
Editor

Computational Acoustics

Springer

Editor
Manfred Kaltenbacher
Institute of Mechanics and Mechatronics
Vienna University of Technology
Vienna
Austria

ISSN 0254-1971 ISSN 2309-3706 (electronic)
CISM International Centre for Mechanical Sciences
ISBN 978-3-319-59037-0 ISBN 978-3-319-59038-7 (eBook)
DOI 10.1007/978-3-319-59038-7

Library of Congress Control Number: 2017941071

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The book is a result of the Advanced School Computational Acoustics, which took place at the International Centre for Mechanical Sciences (CISM), Udine, Italy, in May 2016.

The aim of this book is to present state-of-the-art overview of numerical schemes efficiently solving the acoustic conservation equations, the acoustic wave equation, and its Fourier-transformed Helmholtz equation. Thereby, the different equations model both vibrational and flow-induced sound generation and its propagation.

Chapter "Fundamental Equations of Acoustics" sets the scene by providing the mathematical/physical modeling of acoustic fields. Thereby, the equations of acoustics are based on the general equations of fluid dynamics: conservation of mass, momentum, energy, and closed by the appropriate constitutive equations defining the thermodynamic state. The use of a perturbation ansatz, which decomposes the physical quantities such as density, pressure, and velocity into mean, incompressible fluctuating and compressible fluctuating ones, allows to derive linearized acoustic conservation equations and its state equation. Thereby, we derive acoustic wave equations for both homogeneous and inhomogeneous media.

Chapter "Non-conforming Finite Elements for Flexible Discretization with Applications to Aeroacoustics" focuses toward non-conforming finite elements for flexible discretization. Therewith, we allow for each subdomain an optimal grid. The two proposed methods—Mortar and Nitsche-type mortaring—fulfill the physical conditions along the non-conforming interfaces. We exploit this capability and apply it to real engineering applications in aeroacoustic. The results demonstrate the superiority of the nonconforming finite elements over standard finite elements concerning preprocessing, mesh generation flexibility, accuracy, and computational time.

Chapter "Boundary Element Method for Time–Harmonic Acoustic Problems" presents the solution of time–harmonic acoustic problems by the boundary element method (BEM). Specifically, the Helmholtz equation with admittance boundary conditions is solved in three-dimensional space. The chapter starts with a derivation of the Kirchhoff–Helmholtz integral equation from a residual formulation of the

Helmholtz equation. The discretization process with introduction of basis and test functions is described and shown for the collocation and the Galerkin method. Throughout the chapter, numerous different examples are presented, both simple one-dimensional examples having analytical solutions, which may be used for implementation verification, and rather industrial applications such as sedan cabin compartments, diesel engine radiation, and tire noise problems demonstrating the applicability.

Chapter "Direct Aeroacoustic Simulations Based on High Order Discontinuous Galerkin Schemes" focuses on direct aeroacoustic simulations based on high-order discontinuous Galerkin schemes. The framework presented is based on a particular version of the Discontinuous Galerkin method, in which a nodal as well as discretely orthogonal basis is used for computational efficiency. This discretization choice allows arbitrary order in space while also supporting unstructured meshes. After discussing the details of the framework, examples of direct noise computation are presented, with a special focus on the numerical simulation of acoustic feedback in a complex automotive application.

Numerical schemes lead to a system of algebraic equations, which needs efficient solvers. Therefore, Chapter "Direct and Iterative Solvers" presents a compact introduction to direct and iterative solvers for systems of algebraic equations typically arising from the finite element discretization of partial differential equations. Beside classical iterative solvers, we also consider advanced preconditioning and solving techniques like additive and multiplicative Schwarz methods, generalizing Jacobis and Gauss-Seidel's ideas to more general subspace correction methods. In particular, we consider multilevel diagonal scaling and multigrid methods.

We have pleasure in thanking our colleagues, Gary Cohen, Dan Givoli, Ulrich Langer, Steffen Marburg, Claus-Dieter Munz, and Martin Neumüller for presenting their lectures, and the students for attending the course and contributing to discussions. Furthermore, we particularly thank the Rectors and officers at CISM for their enthusiasm, assistance, and hospitality. Finally, we want to thank Springer for their kind assistance, and especially Sooryadeepth Jayakrishnan and his team for their great job in doing the layout.

Vienna, Austria Manfred Kaltenbacher

Contents

Fundamental Equations of Acoustics

Manfred Kaltenbacher

Abstract The equations of acoustics are based on the general equations of fluid dynamics: conservation of mass, momentum, energy and closed by the appropriate constitutive equation defining the thermodynamic state. The use of a perturbation ansatz, which decomposes the physical quantities density, pressure and velocity into mean, incompressible fluctuating and compressible fluctuating ones, allows to derive linearized acoustic conservation equations and its state equation. Thereby, we derive acoustic wave equations both for homogeneous and inhomogeneous media, and the equations model both vibrational- and flow-induced sound generation and its propagation.

1 Overview

Acoustics has developed into an interdisciplinary field encompassing the disciplines of physics, engineering, speech, audiology, music, architecture, psychology, neuroscience, and others (see, e.g., Rossing 2007). Therewith, the arising multi-field problems range from classical airborne sound over underwater acoustics (e.g., ocean acoustics) to ultrasound used in medical application. Here, we concentrate on the basic equations of acoustics describing acoustic phenomena. Thereby, we start with the mass, momentum and energy conservation equations of fluid dynamics as well as the constitutive equations. Furthermore, we introduce the Helmholtz decomposition to split the overall fluid velocity in a pure solenoidal (incompressible part) and irrotational (compressible) part. Since, wave propagation needs a compressible medium, we associate this part to acoustics. Furthermore, we apply a perturbation method to derive the acoustic wave equation, and discuss the main physical quantities of acoustics, plane and spherical wave solutions. Finally, we focus towards the two main mechanism of sound generation: aeroacoustics (flow induced sound) and vibroacoustics (sound generation due to mechanical vibrations).

M. Kaltenbacher (✉)
Institute of Mechanics and Mechatronics, TU Wien, Vienna, Austria
e-mail: manfred.kaltenbacher@tuwien.ac.at

M. Kaltenbacher (ed.), *Computational Acoustics*, CISM International Centre for Mechanical Sciences 579, DOI 10.1007/978-3-319-59038-7_1

2 Basic Equations of Fluid Dynamics

We consider the motion of fluids in the continuum approximation, so that a body $\mathcal{B}$ is composed of particles $\mathcal{R}$ as displayed in Fig. 1. Thereby, a particle $\mathcal{R}$ already represents a macroscopic element. On the one hand a particle has to be small enough to describe the deformation accurately and on the other hand large enough to satisfy the assumptions of continuum theory. This means that the physical quantities density ρ, pressure p, velocity $\boldsymbol{v}$, and so on are functions of space and time, and are written as density $\rho(x_i, t)$, pressure $p(x_i, t)$, velocity $\boldsymbol{v}(x_i, t)$, etc. So, the total change of a scalar quantity like the density ρ is

$$\mathrm{d}\rho = \left(\frac{\partial \rho}{\partial t}\right) \mathrm{d}t + \left(\frac{\partial \rho}{\partial x_1}\right) \mathrm{d}x_1 + \left(\frac{\partial \rho}{\partial x_2}\right) \mathrm{d}x_2 + \left(\frac{\partial \rho}{\partial x_3}\right) \mathrm{d}x_3 \,. \tag{1}$$

Therefore, the total derivative (also called substantial derivative) computes by

$$\begin{aligned} \frac{\mathrm{d}\rho}{\mathrm{d}t} &= \frac{\partial \rho}{\partial t} + \frac{\partial \rho}{\partial x_1}\left(\frac{\mathrm{d}x_1}{\mathrm{d}t}\right) + \frac{\partial \rho}{\partial x_2}\left(\frac{\mathrm{d}x_2}{\mathrm{d}t}\right) + \frac{\partial \rho}{\partial x_3}\left(\frac{\mathrm{d}x_3}{\mathrm{d}t}\right) \\ &= \frac{\partial \rho}{\partial t} + \sum_{i=1}^{3} \frac{\partial \rho}{\partial x_i}\left(\frac{\mathrm{d}x_i}{\mathrm{d}t}\right) = \frac{\partial \rho}{\partial t} + \frac{\partial \rho}{\partial x_i} \underbrace{\left(\frac{\mathrm{d}x_i}{\mathrm{d}t}\right)}_{v_i} . \end{aligned} \tag{2}$$

Note that in the last line of (2) we have used the summation rule of Einstein.[1] Furthermore, in literature the substantial derivative of a physical quantity is mainly denoted by the capital letter D and for an Eulerian frame of reference writes as

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \boldsymbol{v} \cdot \nabla \,. \tag{3}$$

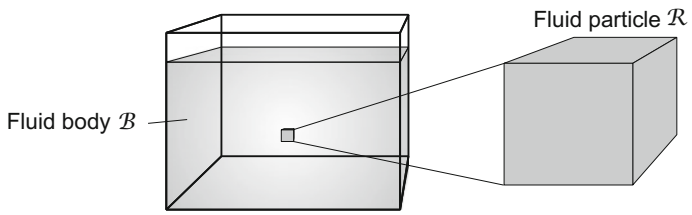


Fig. 1 A body $\mathcal{B}$ composed of particles $\mathcal{R}$

[1] In the following, we will use both vector and index notation; for the main operations see Appendix.

2.1 Spatial Reference Systems

A spatial reference system defines how the motion of a continuum is described i.e., from which perspective an observer views the matter. In a Lagrangian frame of reference, the observer monitors the trajectory in space of each material point and measures its physical quantities. This can be understood by considering a measuring probe which moves together with the material, like a boat on a river. The advantage is that free or moving boundaries can be captured easily as they require no special effort. Therefore, the approach is suitable in the case of structural mechanics. However, its limitation is obtained dealing with large deformation, as in the case of fluid dynamics. In this case, a better choice is the Eulerian frame of reference, in which the observer monitors a single point in space when measuring physical quantities – the measuring probe stays at a fixed position in space. However, contrary to the Lagrangian approach, difficulties arise with deformations on the domain boundary, e.g., free boundaries and moving interfaces.

To derive integral formulations of balance equations, the rate of change of integrals of scalar and vector functions has to be described, which is known as the Reynolds' transport theorem. The volume integral can change for two reasons: (1) scalar or vector functions change (2) the volume changes. The following discussion is directed to scalar valued functions. In an Eulerian context, time derivation must also take the time dependent domain $\Omega(t)$ into account by adding a surface flux term, which can be formulated as a volume term using the integral theorem of Gauß. This results in

$$\begin{aligned} \frac{D}{Dt}\int\limits_{\Omega(t)} f \,\mathrm{d}\boldsymbol{x} &= \int\limits_{\Omega(t)} \frac{\partial}{\partial t} f \,\mathrm{d}\boldsymbol{x} + \int\limits_{\Gamma(t)} f\boldsymbol{v}\cdot\boldsymbol{n}\,\mathrm{d}s \\ &= \int\limits_{\Omega(t)} \left(\frac{\partial}{\partial t} f + \nabla\cdot(f\boldsymbol{v})\right) \mathrm{d}\boldsymbol{x}\,. \end{aligned} \tag{4}$$

2.2 Conservation Equations

The basic equations for the flow field are the conservation of mass, momentum and energy. Together with the constitutive equations and equations of state, a full set of partial differential equations (PDEs) is derived.

Conservation of mass The mass m of a body is the volume integral of its density ρ,

$$m = \int\limits_{\Omega(t)} \rho(\boldsymbol{x}, t)\,\mathrm{d}\boldsymbol{x}\,. \tag{5}$$

Mass conservation states that the mass of a body is conserved over time, assuming there is no source or drain. Therefore, applying Reynolds' transport theorem (4), results in

$$\begin{aligned}\frac{Dm}{Dt} &= \int_\Omega \frac{\partial \rho}{\partial t}\,\mathrm{d}\boldsymbol{x} + \int_\Gamma \rho \boldsymbol{v}\cdot\boldsymbol{n}\,\mathrm{d}s \\ &= \int_\Omega \left(\frac{\partial \rho}{\partial t} + \nabla\cdot(\rho\boldsymbol{v})\right)\mathrm{d}\boldsymbol{x} = 0.\end{aligned} \tag{6}$$

The integral in (6) can be dismissed, as it holds for arbitrary Ω and in the special case of an incompressible fluid ($\rho = const.\quad \forall(\boldsymbol{x},t)\in\Omega\times\mathbb{R}$), which may be assumed for low Mach numbers (see Sect. 2.4), the time and space derivative of the density vanishes. This leads to the following form of mass conservation equations

$$\begin{aligned}\frac{\partial \rho}{\partial t} + \nabla\cdot(\rho\boldsymbol{v}) = 0 \qquad &\text{(compressible fluid)},\\ \nabla\cdot\boldsymbol{v} = 0 \qquad &\text{(incompressible fluid)}.\end{aligned} \tag{7}$$

Conservation of momentum The equation of momentum is implied by Newtons second law and states that momentum $\boldsymbol{I}_\mathrm{m}$ is the product of mass m and velocity $\boldsymbol{v}$

$$\boldsymbol{I}_\mathrm{m} = m\boldsymbol{v}\,. \tag{8}$$

Derivation in time gives the rate of change of momentum, which is equal to the force $\boldsymbol{F}$ and reveals the relation to Newtons second law in an Eulerian reference system

$$\boldsymbol{F} = \frac{D\boldsymbol{I}_\mathrm{m}}{Dt} = \frac{D}{Dt}(m\boldsymbol{v}) = \frac{\partial}{\partial t}(m\boldsymbol{v}) + \nabla\cdot(m\boldsymbol{v}\otimes\boldsymbol{v}), \tag{9}$$

where $\boldsymbol{v}\otimes\boldsymbol{v}$ is a tensor defined by the dyadic product $\otimes$ (see Appendix). The last equality in (9) is derived from Reynolds transport theorem (4) and mass conservation (7).

The forces $\boldsymbol{F}$ acting on fluids can be split up into forces acting on the surface of the body $\boldsymbol{F}_\Gamma$, forces due to momentum of the molecules $D\boldsymbol{I}_\mathrm{m}/Dt$ and external forces $\boldsymbol{F}_{ex}$ (e.g. gravity, electromagnetic forces)

$$\boldsymbol{F} = \boldsymbol{F}_\Gamma + \frac{D}{Dt}\boldsymbol{I}_\mathrm{m} + \boldsymbol{F}_{ex}. \tag{10}$$

Thereby, the surface force computes by

$$\sum_{i=1}^{3}\boldsymbol{F}_{\Gamma_j} = -\sum_{i=1}^{3}\frac{\partial p}{\partial x_j}\Omega\boldsymbol{n}_j = -\Omega\nabla p. \tag{11}$$

and the total change of momentum $\boldsymbol{I}_\mathrm{m}$ by

$$\frac{D}{Dt}\boldsymbol{I}_\mathrm{m} = \Omega\,\nabla\cdot[\boldsymbol{\tau}] \tag{12}$$

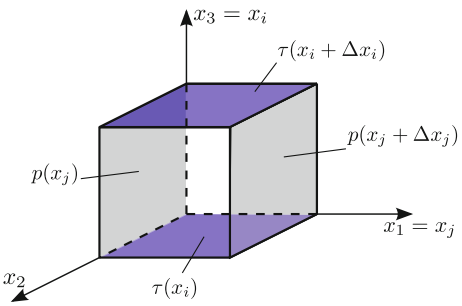


Fig. 2 Forces acting on a fluid element

with the viscous stress tensor $[\boldsymbol{\tau}]$ (see Fig. 2).

Now, we exploit the fact that $m = \rho\Omega$ and insert the pressure force (11), the viscous force (12) and any external forces per unit volume $\boldsymbol{f}$ acting on the fluid into (9). Thereby, we arrive at the momentum equation

$$\frac{\partial \rho \boldsymbol{v}}{\partial t} + \nabla \cdot (\rho\, \boldsymbol{v} \otimes \boldsymbol{v}) = -\nabla p + \nabla \cdot [\boldsymbol{\tau}] + \boldsymbol{f} \tag{13}$$

$$\frac{\partial \rho \boldsymbol{v}}{\partial t} + \nabla \cdot (\rho\, \boldsymbol{v} \otimes \boldsymbol{v} + p\,[\boldsymbol{I}] - [\boldsymbol{\tau}]) = \boldsymbol{f} \tag{14}$$

$$\frac{\partial \rho v_i}{\partial t} + \frac{\partial}{\partial x_j}\left(\rho v_j v_i + p\delta_{ij} - \tau_{ij}\right) = f_i\,, \tag{15}$$

with $[\boldsymbol{I}]$ the identity tensor. Furthermore, we introduce the momentum flux tensor $[\boldsymbol{\pi}]$ defined by

$$\pi_{ij} = \rho v_i v_j + p\delta_{ij} - \tau_{ij}\,, \tag{16}$$

and the fluid stress tensor $[\boldsymbol{\sigma}_{\mathrm{f}}]$ by

$$[\boldsymbol{\sigma}_{\mathrm{f}}] = -p\,[\boldsymbol{I}] + [\boldsymbol{\tau}]\,. \tag{17}$$

To arrive at an alternative formulation for the momentum equation, also called the non-conservative form, we exploit the following identities

$$\nabla \cdot (\rho \boldsymbol{v} \otimes \boldsymbol{v}) = \rho\, \boldsymbol{v} \cdot \nabla \boldsymbol{v} + \boldsymbol{v}\, \nabla \cdot (\rho \boldsymbol{v}) \tag{18}$$

$$\frac{\partial \rho \boldsymbol{v}}{\partial t} = \rho \frac{\partial \boldsymbol{v}}{\partial t} + \boldsymbol{v} \frac{\partial \rho}{\partial t} \tag{19}$$

and rewrite (13) by

$$\rho \frac{\partial \boldsymbol{v}}{\partial t} + \boldsymbol{v}\, \frac{\partial \rho}{\partial t} + \boldsymbol{v} \nabla \cdot (\rho \boldsymbol{v}) + \rho \boldsymbol{v} \cdot \nabla \boldsymbol{v} = -\nabla p + \nabla \cdot [\boldsymbol{\tau}] + \boldsymbol{f}\,. \tag{20}$$

Now, we use the mass conservation and arrive at

$$\rho\frac{\partial \boldsymbol{v}}{\partial t} + \rho\,\boldsymbol{v}\cdot\nabla\boldsymbol{v} = -\nabla p + \nabla\cdot[\boldsymbol{\tau}] + \boldsymbol{f} \qquad (21)$$
$$\rho\frac{\partial v_i}{\partial t} + \rho v_j\frac{\partial v_i}{\partial x_j} = -\frac{\partial p}{\partial x_i} + \frac{\partial \tau_{ij}}{\partial x_j} + f_i\,.$$

Conservation of energy The total balance of energy considers the inner, the kinetic and potential energies of a fluid. Since we do not consider gravity, the total change of energy over time for a fluid element with mass m is given by

$$\frac{\mathrm{D}}{\mathrm{D}t}\left(m\left(\frac{1}{2}v^2+e\right)\right) = m\frac{\mathrm{D}}{\mathrm{D}t}\left(\frac{1}{2}v^2+e\right) + \left(\frac{1}{2}v^2+e\right)\frac{\mathrm{D}m}{\mathrm{D}t} \qquad (22)$$

with e the inner energy and $v^2 = \boldsymbol{v}\cdot\boldsymbol{v}$. Due to mass conservation, the second term is zero and we obtain

$$\frac{\mathrm{D}}{\mathrm{D}t}\left(m\left(\frac{1}{2}v^2+e\right)\right) = \rho\Omega\frac{\mathrm{D}}{\mathrm{D}t}\left(\frac{1}{2}v^2+e\right)\,. \qquad (23)$$

This change of energy can be caused by Durst (2006)

- heat production per unit of volume: $q_\mathrm{h}\,\Omega$
- heat conduction energy due to heat flux $\boldsymbol{q}_\mathrm{T}$: $(-\partial q_{\mathrm{T}i}/\partial x_i)\,\Omega$
- energy due to surface pressure force: $(-\partial/\partial x_i(pv_i))\,\Omega$
- energy due to surface shear force: $\left(-\partial/\partial x_i(\tau_{ij}v_j)\right)\Omega$
- mechanical energy due to the force density $\boldsymbol{f}_i$ given by: $(f_i v_i)\,\Omega$

Thereby, we arrive at the conservation of energy given by

$$\rho\,\frac{\mathrm{D}}{\mathrm{D}t}\left(\frac{1}{2}v^2+e\right) = -\frac{\partial q_{\mathrm{T}i}}{\partial x_i} - \frac{\partial p v_i}{\partial x_i} - \frac{\partial \tau_{ij}v_j}{\partial x_i} + f_i v_i + q_\mathrm{h} \qquad (24)$$

or in vector notation by

$$\rho\,\frac{\mathrm{D}}{\mathrm{D}t}\left(\frac{1}{2}v^2+e\right) = -\nabla\cdot\boldsymbol{q}_\mathrm{T} - \nabla\cdot(p\boldsymbol{v}) - \nabla\cdot([\boldsymbol{\tau}]\cdot\boldsymbol{v}) + \boldsymbol{f}\cdot\boldsymbol{v} + q_\mathrm{h}\,. \qquad (25)$$

By further exploring thermodynamic relations (see Sect. 2.3) and the mechanical energy (obtained by inner product of momentum conservation with $\boldsymbol{v}$), we may write (25) by the specific entropy s as follows Howe (1998)

$$\rho T\frac{\mathrm{D}s}{\mathrm{D}t} = \tau_{ij}\frac{\partial v_i}{\partial x_j} - \frac{\partial q_{\mathrm{T}i}}{\partial x_i} + q_\mathrm{h}\,. \qquad (26)$$

When heat transfer is neglected, the flow ist *adiabatic*. It is *isentropic*, when it is adiabatic and reversible, which means that the viscous dissipation can be neglected, which leads to (no heat production)

$$\rho T \frac{\mathrm{D}s}{\mathrm{D}t} = 0\,. \tag{27}$$

Finally, when the fluid is homogeneous and the entropy uniform ($\mathrm{d}s = 0$), we call the flow *homentropic*.

2.3 Constitutive Equations

The conservation of mass, momentum and energy involve much more unknowns than equations. To close the system, additional information is provided by empirical information in form of constitutive equations. A good approximation is obtained by assuming the fluid to be in thermodynamic equilibrium. This implies for a homogeneous fluid that two *intrinsic* state variables fully determine the state of the fluid.

When we apply specific heat production q_h to a fluid element, then the specific inner energy e increases and at the same time the volume changes by $p\,\mathrm{d}\rho^{-1}$. This thermodynamic relation is expressed by

$$\mathrm{d}e = \mathrm{d}q_\mathrm{h} - p\,\mathrm{d}\rho^{-1}\,, \tag{28}$$

where the second term describes the work done on the fluid element by the pressure. If the change occurs sufficiently slowly, the fluid element is always in thermodynamic equilibrium, and we can express the heat input by the specific entropy s

$$\mathrm{d}q_\mathrm{h} = T\,\mathrm{d}s\,. \tag{29}$$

Therefore, we may rewrite (28) and arrive at the fundamental law of thermodynamics

$$\begin{aligned}\mathrm{d}e &= T\,\mathrm{d}s - p\,\mathrm{d}\rho^{-1}\\ &= T\,\mathrm{d}s + \frac{p}{\rho^2}\,\mathrm{d}\rho\,.\end{aligned} \tag{30}$$

Towards acoustics, it is convenient to choose the mass density ρ and the specific entropy s as intrinsic state variables. Hence, the specific inner energy e is completely defined by a relation denoted as the thermal equation of state

$$e = e(\rho,\, s)\,. \tag{31}$$

Therefore, variations of e are given by

$$de = \left(\frac{\partial e}{\partial \rho}\right)_s \mathrm{d}\rho + \left(\frac{\partial e}{\partial s}\right)_\rho \mathrm{d}s\,. \tag{32}$$

A comparison with the fundamental law of thermodynamics (30) provides the thermodynamic equations for the temperature T and pressure p

$$T = \left(\frac{\partial e}{\partial s}\right)_\rho \; ; \quad p = \rho^2 \left(\frac{\partial e}{\partial \rho}\right)_s . \tag{33}$$

Since p is a function of ρ and s, we may write

$$dp = \left(\frac{\partial p}{\partial \rho}\right)_s \mathrm{d}\rho + \left(\frac{\partial p}{\partial s}\right)_\rho \mathrm{d}s . \tag{34}$$

As sound is defined as isentropic ($ds = 0$) pressure-density perturbations, the isentropic speed of sound is defined by

$$c = \sqrt{\left(\frac{\partial p}{\partial \rho}\right)_s} . \tag{35}$$

Since in many applications the fluid considered is air at ambient pressure and temperature, we may use the ideal gas law

$$p = \rho R T \tag{36}$$

with the specific gas constant R, which computes for an ideal gas as

$$R = c_p - c_\Omega . \tag{37}$$

In (37) c_p, c_Ω denote the specific heat at constant pressure and constant volume, respectively. Furthermore, the inner energy e depends for an ideal gas just on the temperature T via

$$\mathrm{d}e = c_\Omega \, \mathrm{d}T . \tag{38}$$

Substituting this relations in (30), assuming an isentropic state ($\mathrm{d}s = 0$) and using (36) results in

$$c_\Omega \, \mathrm{d}T = \frac{p}{\rho^2} \mathrm{d}\rho \;\rightarrow\; \frac{\mathrm{d}T}{T} = \frac{R}{c_\Omega} \frac{\mathrm{d}\rho}{\rho} . \tag{39}$$

Using (36), the total change $\mathrm{d}p$ normalized to p computes as

$$\frac{\mathrm{d}p}{p} = \frac{\mathrm{d}\rho}{\rho} + \frac{\mathrm{d}T}{T} . \tag{40}$$

This relation and applying (39), (37) leads to

$$\frac{\mathrm{d}p}{p} = \frac{\mathrm{d}\rho}{\rho} + \frac{R}{c_\Omega} \frac{\mathrm{d}\rho}{\rho} = \frac{c_p}{c_\Omega} \frac{\mathrm{d}\rho}{\rho} = \kappa \frac{\mathrm{d}\rho}{\rho} \tag{41}$$

with κ the specific heat ratio (also known as adiabatic exponent). A comparison of (41) with (35) yields

$$c = \sqrt{\kappa p/\rho} = \sqrt{\kappa R T}\,. \tag{42}$$

We see that the speed of sound c of an ideal gas depends only on the temperature. For air κ has a value of 1.402 so that we obtain a speed of sound c at $T = 15\,^{\circ}\mathrm{C}$ of 341 m/s. For most practical applications, we can set the speed of sound to 340 m/s within a temperature range of 5–25 °C. Combining (41) and (42), we obtain the general pressure-density relation for an isotropic state

$$\frac{\mathrm{d}p}{\mathrm{d}t} = c^2 \frac{\mathrm{d}\rho}{\mathrm{d}t}\,. \tag{43}$$

Furthermore, since we use an Eulerian frame of reference, we may rewrite (43) by

$$\frac{Dp}{Dt} = c^2 \frac{D\rho}{Dt}\,. \tag{44}$$

For liquids, such as water, the pressure-density relation is written by the adiabatic bulk modulus K_s (or its reciprocal $1/K_\mathrm{s}$, known as the adiabatic compressibility) and (43) reads as

$$\frac{Dp}{Dt} = \frac{K_\mathrm{s}}{\rho} \frac{D\rho}{Dt}\,. \tag{45}$$

2.4 Characterization of Flows by Dimensionless Numbers

Two flows around geometric similar models are physically similar if all characteristic numbers coincide (Schlichting and Gersten 2006). Especially for measurement setups, these similarity considerations are important as it allows measuring of down sized geometries. Furthermore, the characteristic numbers are used to classify a flow situation. The Reynolds number is defined by

$$\mathrm{Re} = \frac{vl}{\nu} \tag{46}$$

with the characteristic flow velocity v, flow length l and kinematic viscosity ν. It provides the ratio between stationary inertia forces and viscous forces. Thereby, it allows to subdivide flows into laminar and turbulent ones. The Mach number allows for an approximative subdivision of a flow in compressible ($\mathrm{Ma} > 0.3$) and incompressible ($\mathrm{Ma} \leq 0.3$), and is defined by

$$\mathrm{Ma} = \frac{v}{c} \tag{47}$$

with c the speed of sound. In unsteady problems, periodic oscillating flow structures may occur, e.g. the Kármán vortex street in the wake of a cylinder. The dimensionless frequency of such an oscillation is denoted as the Strouhal number, and is defined by

$$\mathrm{St} = f\,\frac{l}{v} \tag{48}$$

with f the shedding frequency.

2.5 Towards Acoustics

According to the Helmholtz decomposition, the velocity vector $\boldsymbol{v}$ (as any vector field) can be split into an irrotational part and a solenoidal part

$$\boldsymbol{v} = \nabla\phi + \nabla \times \boldsymbol{\Psi}\,, \tag{49}$$

where ϕ is a scalar potential and $\boldsymbol{\Psi}$ a vector potential. Thereby, we call a flow being purely described by a scalar potential via

$$\boldsymbol{v} = \nabla\phi$$

a *potential flow*. Using (49), mass conservation (see (7)) may be written as

$$\begin{aligned}\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\boldsymbol{v}) &= \frac{\partial\rho}{\partial t} + \boldsymbol{v}\cdot\nabla\rho + \rho\nabla\cdot\boldsymbol{v}\\ &= \frac{D\rho}{Dt} + \rho\nabla\cdot\nabla\phi + \rho\underbrace{\nabla\cdot\nabla\times\boldsymbol{\Psi}}_{=0}\\ \frac{1}{\rho}\frac{D\rho}{Dt} &= -\nabla\cdot\nabla\phi\,.\end{aligned} \tag{50}$$

This result obviously leads us to the interpretation that the flow related to the acoustic field is an irrotational flow and that the acoustic field is the unsteady component of the gradient of the velocity potential ϕ. On the other hand, taking the curl of (49) results in the vorticity of the flow

$$\boldsymbol{\omega} = \nabla\times\boldsymbol{v} = \nabla\times\nabla\times\boldsymbol{\Psi} + \underbrace{\nabla\times\nabla\phi}_{=0} = \nabla\times\nabla\times\boldsymbol{\Psi}\,. \tag{51}$$

We see that this quantity is fully defined by the vector potential and characterizes the solenoidal part of the flow field.

3 Basic Equations of Acoustics

3.1 Acoustic Wave Equation

We assume an isentropic case, where the total variation of the entropy is zero and the pressure is only a function of the density (see (44)). Furthermore, we restrict ourself to a perfect (non-viscous) fluid (setting the viscous fluid tensor $[\boldsymbol{\tau}]$ to zero) and neglect external force density $\boldsymbol{f}$. Thereby, we arrive, according to Sect. 2, to the following set of equations

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \boldsymbol{v}) = 0 \tag{52}$$

$$\rho \frac{\partial \boldsymbol{v}}{\partial t} + \rho\, \boldsymbol{v} \cdot \nabla \boldsymbol{v} + \nabla p = 0 \tag{53}$$

$$\frac{Dp}{Dt} = c^2 \frac{D\rho}{Dt}\,. \tag{54}$$

In a first step, we consider the static case with mean pressure p_0, mean density ρ_0 and velocity $\boldsymbol{v}_0$ being zero. Therefore, (52) is fulfilled identically, while (53) results in

$$\nabla p_0 = 0\,. \tag{55}$$

Furthermore, (54) is automatically satisfied by some function c_0 (independent of t) defined by means of some virtual non-static variations of the solution. In a next step, we consider a non-static solution of very small order according to a perturbation of the mean quantities

$$p = p_0 + p_\mathrm{a}\,; \quad \rho = \rho_0 + \rho_\mathrm{a}\,; \quad \boldsymbol{v} = \boldsymbol{v}_\mathrm{a} \tag{56}$$

with the following relations

$$p_\mathrm{a} \ll p_0\,; \quad \rho_\mathrm{a} \ll \rho_0\,. \tag{57}$$

We name p_a the acoustic pressure, ρ_a the acoustic density and $\boldsymbol{v}_\mathrm{a}$ the acoustic particle velocity. Using the perturbation ansatz (56) and substituting it into (52)–(54), results in

$$\frac{\partial(\rho_0 + \rho_\mathrm{a})}{\partial t} + \nabla \cdot \Big((\rho_0 + \rho_a)\boldsymbol{v}_\mathrm{a}\Big) = 0 \tag{58}$$

$$(\rho_0 + \rho_\mathrm{a}) \frac{\partial \boldsymbol{v}_\mathrm{a}}{\partial t} + \Big((\rho_0 + \rho_\mathrm{a})\boldsymbol{v}_\mathrm{a}\Big) \cdot \nabla \boldsymbol{v}_\mathrm{a} + \nabla\big(p_0 + p_\mathrm{a}\big) = 0 \tag{59}$$

$$\left(\frac{\partial}{\partial t} + \boldsymbol{v}_\mathrm{a} \cdot \nabla\right)(p_0 + p_\mathrm{a}) - c_0^2 \left(\frac{\partial}{\partial t} + \boldsymbol{v}_\mathrm{a} \cdot \nabla\right)(\rho_0 + \rho_\mathrm{a}) = 0\,. \tag{60}$$

In a next step, we are allowed to cancel second order terms (e.g., such as $\rho_\mathrm{a}\boldsymbol{v}_\mathrm{a}$), consider that p_0 does not vary over space (see (55)) and arrive at

$$\frac{\partial \rho_\mathrm{a}}{\partial t} + \nabla(\rho_0 \boldsymbol{v}_\mathrm{a}) = q_\mathrm{ma} \tag{61}$$

$$\rho_0 \frac{\partial \boldsymbol{v}_\mathrm{a}}{\partial t} + \nabla p_\mathrm{a} = \boldsymbol{q}_\mathrm{mo} \tag{62}$$

$$\frac{\partial p_\mathrm{a}}{\partial t} = c_0^2 \left(\frac{\partial \rho_\mathrm{a}}{\partial t} + \boldsymbol{v}_\mathrm{a} \cdot \nabla \rho_0 \right) . \tag{63}$$

Here, we have included possible modeled source terms in (61) (linearized conservation of mass) and (62) (linearized conservation of momentum). Please note that just in the case of constant mean density, i.e. $\nabla \rho_0 = 0$, we are allowed to express the acoustic pressure-density relation by

$$p_\mathrm{a} = c_0^2 \rho_\mathrm{a} \,, \tag{64}$$

Now, we use (61), substitute it into (63) and obtain the final two equations for linear acoustics

$$\frac{1}{\rho_0 c_0^2} \frac{\partial p_\mathrm{a}}{\partial t} + \nabla \cdot \boldsymbol{v}_\mathrm{a} = \frac{1}{\rho_0} q_\mathrm{ma} \tag{65}$$

$$\frac{\partial \boldsymbol{v}_\mathrm{a}}{\partial t} + \frac{1}{\rho_0} \nabla p_\mathrm{a} = \frac{1}{\rho_0} \boldsymbol{q}_\mathrm{mo} \,. \tag{66}$$

Applying $\partial/\partial t$ to (65), $\nabla\cdot$ to (66) and subtracting the resulting equations provides the linear wave equation for an inhomogeneous medium (density depending on space)

$$\frac{1}{\rho_0 c_0^2} \frac{\partial^2 p_\mathrm{a}}{\partial t^2} - \nabla \cdot \frac{1}{\rho_0} \nabla p_\mathrm{a} = \frac{1}{\rho_0} \frac{\partial q_\mathrm{ma}}{\partial t} - \nabla \cdot \frac{\boldsymbol{q}_\mathrm{mo}}{\rho_0} \,. \tag{67}$$

Furthermore, since the term $\rho_0 c_0^2$ is constant in space and time, we may rewrite (67) by

$$\frac{\partial^2 p_\mathrm{a}}{\partial t^2} - \nabla \cdot c_0^2 \nabla p_\mathrm{a} = c_0^2 \frac{\partial q_\mathrm{ma}}{\partial t} - \nabla \cdot \left(c_0^2 \boldsymbol{q}_\mathrm{mo} \right) \,. \tag{68}$$

This form of wave equation is mainly used when considering the influence of temperature gradient (speed of sound c_0 depends on temperature, see (42)) on wave propagation. For liquids, (67) may be written as

$$\frac{1}{K_\mathrm{s}} \frac{\partial^2 p_\mathrm{a}}{\partial t^2} - \nabla \cdot \frac{1}{\rho_0} \nabla p_\mathrm{a} = \frac{1}{\rho_0} \frac{\partial q_\mathrm{ma}}{\partial t} - \nabla \cdot \frac{\boldsymbol{q}_\mathrm{mo}}{\rho_0} \,. \tag{69}$$

By applying the chain rule

$$\nabla \cdot \frac{1}{\rho_0} \nabla p_\mathrm{a} = \frac{1}{\rho_0} \nabla \cdot \nabla p_\mathrm{a} - \frac{1}{\rho_0^2} \nabla \rho_0 \cdot \nabla p_\mathrm{a} \,,$$

we arrive at

$$\frac{1}{K_\mathrm{s}} \frac{\partial^2 p_\mathrm{a}}{\partial t^2} - \frac{1}{\rho_0} \nabla \cdot \nabla p_\mathrm{a} + \frac{1}{\rho_0^2} \nabla \rho_0 \cdot \nabla p_\mathrm{a} = \frac{1}{\rho_0} \frac{\partial q_\mathrm{ma}}{\partial t} - \nabla \cdot \frac{\boldsymbol{q}_\mathrm{mo}}{\rho_0} \,. \tag{70}$$

This form of the wave equation explicitly shows the influence of a space dependent density ρ_0.

A wave equation for the particle velocity $\boldsymbol{v}_\mathrm{a}$ may be derived by rewriting (65), (66) as

$$\frac{\partial p_\mathrm{a}}{\partial t} + \rho_0 c_0^2 \nabla \cdot \boldsymbol{v}_\mathrm{a} = c_0^2 \, q_\mathrm{ma} \tag{71}$$

$$\rho_0 \frac{\partial \boldsymbol{v}_\mathrm{a}}{\partial t} + \nabla p_\mathrm{a} = \boldsymbol{q}_\mathrm{mo} \,. \tag{72}$$

Now, we apply ∇ to (71), $\partial/\partial t$ to (72) and by subtract the resulting equations we arrive at

$$\rho_0 \frac{\partial^2 \boldsymbol{v}_\mathrm{a}}{\partial t^2} - \nabla \rho_0 c_0^2 \nabla \cdot \boldsymbol{v}_\mathrm{a} = \frac{\partial \boldsymbol{q}_\mathrm{mo}}{\partial t} - \nabla c_0^2 q_\mathrm{ma} \,. \tag{73}$$

It is a vectorial wave equation coupling the three components of the particle velocity. Since the particle velocity $\boldsymbol{v}_\mathrm{a}$ is irrotational, we may express it by the scalar acoustic potential ψ_a via

$$\boldsymbol{v}_\mathrm{a} = -\nabla \psi_\mathrm{a} \,. \tag{74}$$

Substituting this relation into (73), assuming zero source terms and constant density condition ($\nabla \rho_0 = 0$) results in

$$\nabla \left(\frac{\partial^2 \psi_\mathrm{a}}{\partial t^2} - c_0^2 \nabla \cdot \nabla \psi_\mathrm{a} \right) = 0 \,. \tag{75}$$

This equation is clearly satisfied, when ψ_a fulfills

$$\frac{1}{c_0^2} \frac{\partial^2 \psi_\mathrm{a}}{\partial t^2} - \nabla \cdot \nabla \psi_\mathrm{a} = 0 \,. \tag{76}$$

Finally, we provide the most used wave equation in terms of the acoustic pressure p_a, which however does not model an inhomogeneous fluid. It is obtained from (67) assuming a space constant speed of sound c_0 and mean density ρ_0

$$\frac{1}{c_0^2} \frac{\partial^2 p_\mathrm{a}}{\partial t^2} - \nabla \cdot \nabla p_\mathrm{a} = \frac{\partial q_\mathrm{ma}}{\partial t} - \nabla \cdot \boldsymbol{q}_\mathrm{mo} \,. \tag{77}$$

By performing a Fourier transform, we arrive at Helmholtz equation

$$\nabla \cdot \nabla \hat{p}_{\mathrm{a}} + k^2 \hat{p}_{\mathrm{a}} = -j\omega \hat{q}_{\mathrm{ma}} + \nabla \cdot \hat{\boldsymbol{q}}_{\mathrm{mo}} \tag{78}$$

with the Fourier-transformed acoustic pressure $\hat{p}_{\mathrm{a}}$ and source terms $\hat{q}_{\mathrm{ma}}$, $\hat{\boldsymbol{q}}_{\mathrm{mo}}$ as well as angular frequency ω, wave number k and imaginary unit j (see (86)).

3.2 Simple Solutions

In order to get some physical insight in the propagation of acoustic sound, we will consider two special cases: plane and spherical waves. Let's start with the simpler case, the propagation of a plane wave as displayed in Fig. 3. Thus, we can express the acoustic pressure by $p_{\mathrm{a}} = p_{\mathrm{a}}(x,t)$ and the particle velocity by $\boldsymbol{v}_{\mathrm{a}} = v_{\mathrm{a}}(x,t)\boldsymbol{e}_x$. Using these relations together with the linear pressure-density law (assuming constant mean density, see (64)), we arrive at the following 1D linear wave equation

$$\frac{\partial^2 p_{\mathrm{a}}}{\partial x^2} - \frac{1}{c_0^2}\frac{\partial^2 p_{\mathrm{a}}}{\partial t^2} = 0\,, \tag{79}$$

which can be rewritten in factorized version as

$$\left(\frac{\partial}{\partial x} - \frac{1}{c_0}\frac{\partial}{\partial t}\right)\left(\frac{\partial}{\partial x} + \frac{1}{c_0}\frac{\partial}{\partial t}\right) p_{\mathrm{a}} = 0\,. \tag{80}$$

This version of the linearized, 1D wave equation motivates us to introduce the following two functions (solution according to d'Alembert)

$$\xi = t - x/c_0\,; \quad \eta = t + x/c_0$$

with properties

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial \xi} + \frac{\partial}{\partial \eta}\,; \qquad \frac{\partial}{\partial x} = \frac{1}{c_0}\left(\frac{\partial}{\partial \eta} - \frac{\partial}{\partial \xi}\right).$$

Therewith, we obtain for the factorized operator

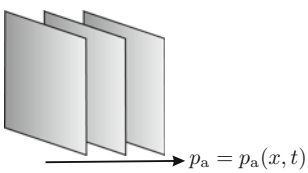


Fig. 3 Propagation of a plane wave

$$\frac{\partial}{\partial x} - \frac{1}{c_0}\frac{\partial}{\partial t} = -\frac{2}{c_0}\frac{\partial}{\partial \xi} \qquad \frac{\partial}{\partial x} + \frac{1}{c_0}\frac{\partial}{\partial t} = \frac{2}{c_0}\frac{\partial}{\partial \eta}$$

and the linear, 1D wave equation transfers to

$$-\frac{4}{c_0^2}\frac{\partial}{\partial \xi}\frac{\partial}{\partial \eta} p_\mathrm{a} = 0\,.$$

The general solution computes as a superposition of arbitrary functions of ξ and η

$$p_\mathrm{a} = f(\xi) + f(\eta) = f(t - x/c_0) + g(t + x/c_0)\,. \tag{81}$$

This solution describes waves moving with the speed of sound c_0 in $+x$ and $-x$ direction, respectively. In a next step, we use the linearized conservation of momentum according to (62), and rewrite it for the 1D case (assuming zero source term)

$$\rho_0 \frac{\partial v_\mathrm{a}}{\partial t} + \frac{\partial p_\mathrm{a}}{\partial x} = 0\,. \tag{82}$$

Now, we just consider a forward propagating wave, i.e. $g(t) = 0$, substitute (81) into (82) and obtain

$$\begin{aligned} v_\mathrm{a} &= -\frac{1}{\rho_0}\int \frac{\partial p_\mathrm{a}}{\partial x}\,\mathrm{d}t = \frac{1}{\rho_0 c_0}\int \frac{\partial f(t - x/c_0)}{\partial t}\,\mathrm{d}t \\ &= \frac{1}{\rho_0 c_0} f(t - x/c_0) = \frac{p_\mathrm{a}}{\rho_0 c_0}\,. \end{aligned} \tag{83}$$

Therewith, the value of the acoustic pressure over acoustic particle velocity for a plane wave is constant. To allow for a general orientation of the coordinate system, a free field plane wave may be expressed by

$$p_\mathrm{a} = f(\boldsymbol{n}\cdot\boldsymbol{x} - c_0 t)\,; \quad \boldsymbol{v}_\mathrm{a} = \frac{\boldsymbol{n}}{\rho_0 c_0} f(\boldsymbol{n}\cdot\boldsymbol{x} - c_0 t)\,, \tag{84}$$

where the direction of propagation is given by the unit vector $\boldsymbol{n}$. A time harmonic plane wave of angular frequency $\omega = 2\pi f$ is usually written as

$$p_\mathrm{a}\,,\ \boldsymbol{v}_\mathrm{a} \sim e^{j(\omega t - \boldsymbol{k}\cdot\boldsymbol{x})} \tag{85}$$

with the wave number (also called wave vector) $\boldsymbol{k}$, which computes by

$$\boldsymbol{k} = k\boldsymbol{n} = \frac{\omega}{c_0}\boldsymbol{n}\,. \tag{86}$$

The second case of investigation will be a spherical wave, where we assume a point source located at the origin. In the first step, we rewrite the linearized wave equation in spherical coordinates and consider that the pressure p_a will just depend on the radius r. Therewith, the Laplace-operator reads as

$$\nabla \cdot \nabla p_\mathrm{a}(r,t) = \frac{\partial^2 p_\mathrm{a}}{\partial r^2} + \frac{2}{r}\frac{\partial p_\mathrm{a}}{\partial r} = \frac{1}{r}\frac{\partial^2 r p_\mathrm{a}}{\partial r^2}$$

and we obtain

$$\frac{1}{r}\frac{\partial^2 r p_\mathrm{a}}{\partial r^2} - \frac{1}{c_0^2}\underbrace{\frac{\partial^2 p_\mathrm{a}}{\partial t^2}}_{\frac{1}{r}\frac{\partial^2 r p_\mathrm{a}}{\partial t^2}} = 0\,. \tag{87}$$

A multiplication of (87) with r results in the same wave equation as obtained for the plane case (see (79)), just instead of p_a we have $r p_\mathrm{a}$. Therefore, the solution of (87) reads as

$$p_\mathrm{a}(r,t) = \frac{1}{r}\left(f(t - r/c_0) + g(t + r/c_0)\right)\,, \tag{88}$$

which means that the pressure amplitude will decrease according to the distance r from the source.

3.3 Acoustic Quantities and Order of Magnitudes

Let us consider a loudspeaker generating sound at a fixed frequency f and a number of microphones recording the sound as displayed in Fig. 4. In a first step, we measure the sound with one microphone fixed at $\boldsymbol{x}_0$, and we will obtain a periodic signal in time with the same frequency f and period time $T = 1/f$. In a second step, we use all microphones and record the pressure at a fixed time t_0. Drawing the obtained values along the individual positions of the microphone, e.g. along the coordinate x, we again obtain a periodic signal, which is now periodic in space. This periodicity is characterized by the wavelength λ and is uniquely defined by the frequency f and the speed of sound c_0 via the relation

$$\lambda = \frac{c_0}{f}\,. \tag{89}$$

Assuming a frequency of 1 kHz, the wavelength in air takes on the value of 0.343 m ($c_0 = 343$ m/s).

Strictly speaking, each acoustic wave has to be considered as transient, having a beginning and an end. However, for some long duration sound, we speak of continuous wave (cw) propagation and we define for the acoustic pressure p_a a mean square pressure $(p_\mathrm{a})^2_\mathrm{av}$ as well as a root mean squared (rms) pressure $p_\mathrm{a,rms}$

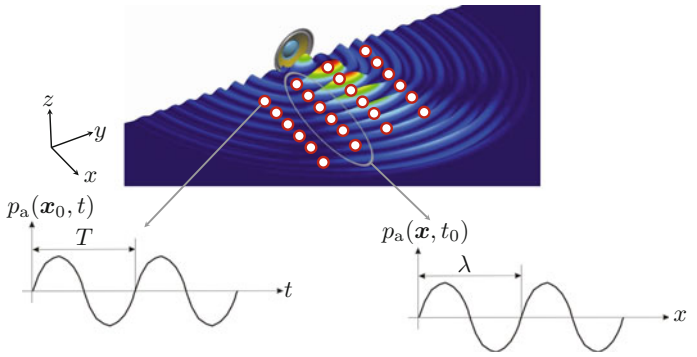


Fig. 4 Sound generated by a loudspeaker and measured by microphones

$$p_{a,rms} = \sqrt{\frac{1}{T}\int_{t_0}^{t_0+T} (p - p_0)^2 \, \mathrm{d}t} = \sqrt{\frac{1}{T}\int_{t_0}^{t_0+T} p_a^2 \, \mathrm{d}t} \, . \tag{90}$$

In (90) T denotes the period time of the signal or if we cannot strictly speak of a periodic signal, an interminable long time interval. Now, it has to be mentioned that the threshold of hearing of an average human is at about 20 μPa and the threshold of pain at about 20 Pa, which differs 10^6 orders of magnitude. Thus, logarithmic scales are mainly used for acoustic quantities. The most common one is the *decibel* (dB), which expresses the quantity as a ratio relative to a reference value. Thereby, the sound pressure level L_{p_a} (SPL) is defined by

$$L_{p_a} = 20 \, \log_{10} \frac{p_{a,\,rms}}{p_{a,ref}} \qquad p_{a,\,ref} = 20 \, \mu \mathrm{Pa} \, . \tag{91}$$

The reference pressure $p_{a,ref}$ corresponds to the sound at 1 kHz that an average person can just hear.

In addition, the acoustic intensity $\boldsymbol{I}_a$ is defined by the product of the acoustic pressure and particle velocity

$$\boldsymbol{I}_a = p_a \boldsymbol{v}_a \, . \tag{92}$$

The intensity level L_{I_a} is then defined by

$$L_{I_a} = 10 \, \log_{10} \frac{I_a^{av}}{I_{a,ref}} \qquad I_{a,ref} = 10^{-12} \, \mathrm{W/m^2} \, , \tag{93}$$

with $I_{a,ref}$ the reference sound intensity corresponding to $p_{a,ref}$. Again, we use an averaged value for defining the intensity level, which computes by

$$I_{\mathrm{a}}^{\mathrm{av}} = \left| \boldsymbol{I}_{\mathrm{a}}^{\mathrm{av}} \right| = \left| \frac{1}{T} \int_{t_0}^{t_0+T} \boldsymbol{v}_{\mathrm{a}} p_{\mathrm{a}} \, \mathrm{d}t \, . \right| \tag{94}$$

Finally, we compute the acoustic power by integrating the acoustic intensity (unit $\mathrm{W/m^2}$) over a closed surface

$$P_{\mathrm{a}} = \oint_{\Gamma} \boldsymbol{I}_{\mathrm{a}} \cdot \mathrm{d}\boldsymbol{s} = \oint_{\Gamma} \boldsymbol{I}_{\mathrm{a}} \cdot \boldsymbol{n} \, \mathrm{d}s \, . \tag{95}$$

Then, the sound-power level $L_{\mathrm{P_a}}$ computes as

$$L_{\mathrm{P_a}} = 10 \, \log_{10} \frac{P_{\mathrm{a}}^{\mathrm{av}}}{P_{\mathrm{a,ref}}} \qquad P_{\mathrm{a,ref}} = 10^{-12} \, \mathrm{W} \, , \tag{96}$$

with $P_{\mathrm{a,ref}}$ the reference sound power corresponding to $p_{\mathrm{a,ref}}$. In Tables 1 and 2 some typical sound pressure and sound power levels are listed.

A useful quantity in acoustics is impedance, which is a measure of the amount by which the motion induced by a pressure applied to a surface is impeded. However, a quantity that varies with time and depends on initial values is not of interest. Thus the impedance is defined via the Fourier transform by

$$\hat{Z}_{\mathrm{a}}(\boldsymbol{x}, \omega) = \frac{\hat{p}_{\mathrm{a}}(\boldsymbol{x}, \omega)}{\hat{\boldsymbol{v}}_{\mathrm{a}}(\boldsymbol{x}, \omega) \cdot \boldsymbol{n}(\boldsymbol{x})} \tag{97}$$

at a point $\boldsymbol{x}$ on the surface Γ with unit normal vector $\boldsymbol{n}$. It is in general a complex number and its real part is called *resistance*, its imaginary part *reactance* and its inverse the *admittance* denoted by $\hat{Y}_{\mathrm{a}}(\boldsymbol{x}, \omega)$. For a plane wave (see Sect. 3.2) the acoustic impedance $\hat{Z}_{\mathrm{a}}$ is constant

$$\hat{Z}_{\mathrm{a}}(\boldsymbol{x}, \omega) = \rho_0 c_0 \, . \tag{98}$$

Table 1 Typical sound pressure levels SPL

Threshold of hearing	Voice at 5 m	Car at 20 m	Pneumatic hammer at 2 m	Jet at 3 m
0 dB	60 dB	80 dB	100 dB	140 dB

Table 2 Typical sound power levels and in parentheses the absolute acoustic power P_{a}

Voice	Fan	Loudspeaker	Jet airliner
30 dB (25 μW)	110 dB (0.05 W)	128 dB (60 W)	170 dB (50 kW)

Often the acoustic impedance $\hat{Z}_\mathrm{a}$ is normalized to this value and then named *specific* impedance (is a dimensionless value).

For a quiescent fluid the acoustic power across a surface Γ computes for time harmonic fields by

$$\begin{aligned} P_\mathrm{a}^{av} &= \int_\Gamma \left(\frac{1}{T} \int_0^T \mathrm{Re}\left(\hat{p}_\mathrm{a} e^{j\omega t}\right) \mathrm{Re}\left(\hat{\boldsymbol{v}}_\mathrm{a} \cdot \boldsymbol{n} e^{j\omega t}\right) \, \mathrm{d}t \right) \mathrm{d}s \\ &= \frac{1}{4} \int_\Gamma \left(\hat{p}_\mathrm{a} \hat{\boldsymbol{v}}_\mathrm{a}^* + \hat{p}_\mathrm{a}^* \hat{\boldsymbol{v}}_\mathrm{a} \right) \cdot \boldsymbol{n} \, \mathrm{d}s \\ &= \frac{1}{2} \int_\Gamma \mathrm{Re}\left(\hat{p}_\mathrm{a}^* \hat{\boldsymbol{v}}_\mathrm{a} \right) \cdot \boldsymbol{n} \, \mathrm{d}s \end{aligned} \tag{99}$$

with * denoting the conjugate complex. Now, we use the impedance $\hat{Z}$ of the surface and arrive at

$$P_\mathrm{a}^{av} = \frac{1}{2} \int_\Gamma \mathrm{Re}\left(\hat{Z}_\mathrm{a}\right) |\hat{\boldsymbol{v}}_\mathrm{a} \cdot \boldsymbol{n}|^2 \, \mathrm{d}s \, . \tag{100}$$

Hence, the real part of the impedance (equal to the resistance) is related to the energy flow. If $\mathrm{Re}\left(\hat{Z}_\mathrm{a}\right) > 0$ the surface is *passive* and absorbs energy, and if $\mathrm{Re}\left(\hat{Z}_\mathrm{a}\right) < 0$ the surface is *active* and produces energy.

In a next step, we analyze what happens, when an acoustic wave propagates from one fluid medium to another one. For simplicity, we restrict to a plane wave, which is described by (see (81))

$$p_\mathrm{a}(t) = f/t - x/c_0) + g(t + x/c_0) \tag{101}$$

In the frequency domain, we may write

$$\hat{p}_\mathrm{a} = \hat{f} e^{-j\omega x/c_0} + \hat{g} e^{j\omega x/c_0} = p^+ e^{j\omega t - jkx} + p^- e^{j\omega t + jkx} \, . \tag{102}$$

Thereby, p^+ is the amplitude of the wave incident at $x = 0$ from $x < 0$ and p^- the amplitude of the reflected wave at $x = 0$ by an impedance $\hat{Z}_\mathrm{a}$. Using the linear conservation of momentum, we obtain the particle velocity

$$\hat{v}_\mathrm{a}(x) = \frac{1}{\rho_0 c_0} \left(p^+ e^{-jkx} - p^- e^{jkx} \right) \, . \tag{103}$$

Defining the reflection coefficient R by

$$R = \frac{p^-}{p^+} \, , \tag{104}$$

we arrive with $\hat{Z}_\mathrm{a} = \hat{p}(0)/\hat{v}(0)$ at

$$R = \frac{\hat{Z}_\mathrm{a} - \rho_0 c_0}{\hat{Z}_\mathrm{a} + \rho_0 c_0} \,. \tag{105}$$

In two dimensions, we consider a plane wave with direction $(\cos\theta, \sin\theta)$, where θ is the angle with the y-axis and the wave approaches from $y < 0$ and hits an impedance $\hat{Z}_\mathrm{a}$ at $y = 0$. The overall pressure may be expressed by

$$\hat{p}_\mathrm{a}(x, y) = e^{-jkx\sin\theta} \left(p^+ e^{-ky\cos\theta} + p^- e^{jky\cos\theta} \right) . \tag{106}$$

Furthermore, the y-component of the particle velocity computes to

$$\hat{v}_\mathrm{a}(x, y) = \frac{\cos\theta}{\rho_0 c_0} e^{-jkx\sin\theta} \left(p^+ e^{-ky\cos\theta} - p^- e^{jky\cos\theta} \right) . \tag{107}$$

Thereby, the impedance is

$$\hat{Z}_\mathrm{a} = \frac{\hat{p}(x, 0)}{\hat{v}(x, 0)} = \frac{\rho_0 c_0}{\cos\theta} \frac{p^+ + p^-}{p^+ - p^-} = \frac{\rho_0 c_0}{\cos\theta} \frac{1 + R}{1 - R} \tag{108}$$

so that the reflection coefficient computes as

$$R = \frac{\hat{Z}_\mathrm{a}\cos\theta - \rho_0 c_0}{\hat{Z}_\mathrm{a}\cos\theta + \rho_0 c_0} \,. \tag{109}$$

4 Boundary Conditions

For realistic simulations, a good approximation of the actual physical boundary conditions is essential. In the two simple cases - acoustically hard and soft boundary - the solution is easy:

- Acoustically hard boundary: Here, the reflection coefficient R gets 1 (total reflection), which means that the surface impedance has to approach infinity. According to (97), the term $\boldsymbol{n} \cdot \boldsymbol{v}_\mathrm{a}$ has to be zero. Using the linearized momentum equation (62) with zero source term, we arrive at the Neumann boundary condition

$$\boldsymbol{n} \cdot \nabla p_\mathrm{a} = \frac{\partial p_\mathrm{a}}{\partial \boldsymbol{n}} = 0 \,. \tag{110}$$

- Acoustically soft boundary: In this case, the acoustic impedance gets zero, which simply results in a homogeneous Dirichlet boundary condition

$$p_{\mathrm{a}} = 0\,. \tag{111}$$

Since real surfaces (boundaries) are never totally hard or totally soft, it seems to be a good idea to use a Robin boundary condition as a model

$$\frac{\partial p_{\mathrm{a}}}{\partial \boldsymbol{n}} + \alpha\, p_{\mathrm{a}} = 0\,. \tag{112}$$

In the time harmonic case, we can explore (62) with zero source term and apply a dot product with the normal vector $\boldsymbol{n}$

$$j\rho_0\omega \boldsymbol{n}\cdot\hat{\boldsymbol{v}}_{\mathrm{a}} + \frac{\partial \hat{p}_{\mathrm{a}}}{\partial \boldsymbol{n}} = 0 \tag{113}$$

By using (97) we obtain

$$\frac{\partial \hat{p}_{\mathrm{a}}}{\partial \boldsymbol{n}} + j\rho_0\omega\frac{\hat{p}_{\mathrm{a}}}{\hat{Z}_{\mathrm{a}}} = 0 \tag{114}$$

and identify the parameter α as

$$\alpha = \frac{j\rho_0\omega}{\hat{Z}_{\mathrm{a}}} = j\rho_0\omega\,\hat{Y}_{\mathrm{a}}\,. \tag{115}$$

As known from measurements, $\hat{Z}_{\mathrm{a}}$ is a function of frequency and therefore a inverse Fourier transform to arrive at a time domain formulation results in a convolution integral. Furthermore, $\hat{Z}_{\mathrm{a}}$ depends on the incident angle of the acoustic wave, which makes acoustic computations of rooms quite complicated. Therefore, often the computational domain is not limited by an impedance boundary condition, but the surrounded elastic body is taken into account (see Sect. 6).

One of the great challenges for wave propagation is the efficient and stable computation of waves in unbounded domains. The crucial point for these computations is that the numerical scheme avoids any reflections at the boundaries, even in case the diameter of the computational domain is just a fraction of a wavelength. Since the eighties of the last century, several numerical techniques have been developed to deal with this topic: infinite elements, Dirichlet-to-Neumann operators based on truncated Fourier expansions, absorbing boundary conditions, etc. The advantages and drawbacks of these different approaches have been widely discussed in literature, see e.g. (Ihlenburg 1998; Givoli 2008). Especially higher order absorbing boundary conditions (ABCs) have gained increasing interest, since these methods do not involve high order derivatives (Hagstrom and Warburton 2009; Bécache et al. 2010).

An alternative approach to approximate free radiation is to surround the computational domain by an additional damping layer and guarantee within the formulation, that no reflections occur at its interface with the computational domain. This so-called perfectly matched layer (PML) technique was first introduced by Berenger (1994) using a splitting of the physical variables and considering a system of first

order partial differential equations (PDEs) for electromagnetics. In the framework of time-harmonic wave propagation, the PML can be interpreted as a complex-valued coordinate stretching (Teixeira and Chew 2000).

5 Aeroacoustics

The sound generated by a flow in an unbounded fluid is usually called *aerodynamic sound*. Most unsteady flows in technical applications are of high Reynolds number, and the acoustic radiation is a very small by-product of the motion. Thereby, the turbulence is usually produced by fluid motion over a solid body and/or by flow instabilities.

Since the beginning of aeroacoustics several numerical methodologies have been proposed. Each of these trying to overcome the challenges that the specific problems pose for an effective and accurate computation of the radiated sound. The main difficulties include (Hardin and Hussaini 1992a, b):

- *Energy disparity and acoustic inefficiency*: There is a large disparity between the overall energy of the flow and the part which is converted to acoustic energy (see Fig. 5). In general, the total radiated power of a turbulent jet scales with $O(v^8/c^5)$, and for a dipole source arising from pressure fluctuations on surfaces inside the flow scales with $O(v^6/c^3)$, where v denotes the characteristic flow velocity and c the speed of sound.
- *Length scale disparity*: A large disparity also occurs between the size of an eddy in the turbulent flow and the wavelength of the generated acoustic sound (see Fig. 5). Low Mach number eddies have a characteristic length scale l and velocity v. This eddy will then radiate acoustic waves of the same characteristic frequency, but with a much larger length scale, expressed by the acoustic wavelength λ

$$\lambda \propto c\frac{l}{v} = \frac{l}{M}\,.$$

- *Simulation of unbounded domains*: As a main issue for the simulation of unbounded domains using volume discretization methods remains the boundary treatment which needs to be applied to avoid the reflection of the outgoing waves on the truncating boundary of the computational domain (see Sect. 4).

Currently, available aeroacoustic methodologies overcome only some of these broad range of numerical and physical issues, which restricts their applicability, making them, in many cases, problem dependent methodologies. In a Direct Numerical Simulation (DNS), all relevant scales of turbulence are resolved and no turbulence modeling is employed. The application of DNS is becoming more feasible with the permanent advancement in computational resources. However, due to the large disparities of length and time scales between fluid and acoustic fields, DNS remains restricted to low Reynolds number flows. Therefore, although some promising work

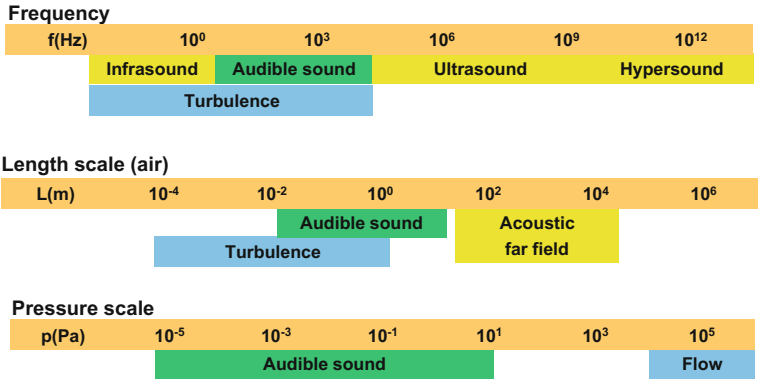
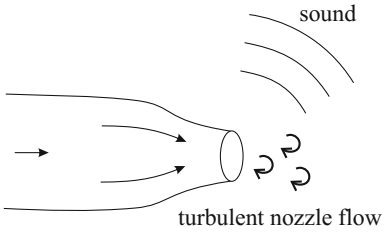


Fig. 5 Flow and acoustic scales

Fig. 6 Turbulent nozzle flow



has been done in this direction (Freund et al. 2000), the simulation of practical problems involving high Reynolds numbers requires very high resolutions and capabilities of supercomputers (Dumbser and Munz 2005; Frank and Munz 2016). Hence, hybrid methodologies have been established as the most practical methods for aeroacoustic computations, due to the separate treatment of the fluid and the acoustic computations. In these schemes, the computational domain is split into a nonlinear source region and a wave propagation region, and different numerical schemes are used for the flow and acoustic computations. Herewith, first a turbulence model is used to compute the unsteady flow in the source region. Secondly, from the fluid field, acoustic sources are evaluated which are then used as input for the computation of the acoustic propagation. In these coupled simulations it is generally assumed that no significant physical effects occur from the acoustic to the fluid field.

5.1 Lighthill's Acoustic Analogy

Lighthill was initially interested in solving the problem, illustrated in Fig. 6, of the sound produced by a turbulent nozzle and arrived at the inhomogeneous wave equation (Lighthill 1952; 1954). For the derivation, we start at Reynolds form of the momentum equation, as given by (15) neglecting any force density $\boldsymbol{f}$

$$\frac{\partial \rho \boldsymbol{v}}{\partial t} + \nabla \cdot [\boldsymbol{\pi}] = 0\,, \tag{116}$$

with the momentum flux tensor $\pi_{ij} = \rho v_i v_j + (p - p_0)\delta_{ij} - \tau_{ij}$, where the constant pressure p_0 is inserted for convenience. In an ideal, linear acoustic medium, the momentum flux tensor contains only the pressure

$$\pi_{ij} \to \pi^0_{ij} = (p - p_0)\delta_{ij} = c_0^2(\rho - \rho_0)\delta_{ij} \tag{117}$$

and Reynolds momentum equation reduces to

$$\frac{\partial \rho v_i}{\partial t} + \frac{\partial}{\partial x_i}\left(c_0^2(\rho - \rho_0)\right) = 0\,. \tag{118}$$

Rewriting the conservation of mass in the form

$$\frac{\partial}{\partial t}(\rho - \rho_0) + \frac{\partial \rho v_i}{\partial x_i} = 0 \tag{119}$$

allows us to eliminate the momentum density ρv_i in (118). Therefore, we perform a time derivative on (119), a spatial derivative on (118) and subtract the two resulting equations. These operations leads to the equation of linear acoustics satisfied by the perturbation density

$$\left(\frac{1}{c_0^2}\frac{\partial^2}{\partial t^2} - \nabla \cdot \nabla\right)\left(c_0^2(\rho - \rho_0)\right) = 0\,. \tag{120}$$

Because flow is neglected, the unique solution of this equation satisfying the radiation condition is $\rho - \rho_0 = 0$.

Now, it can be asserted that the sound generated by the turbulence in the *real fluid* is exactly equivalent to that produced in the ideal, stationary acoustic medium forced by the stress distribution

$$L_{ij} = \pi_{ij} - \pi^0_{ij} = \rho v_i v_j + \left((p - p_0) - c_0^2(\rho - \rho_0)\right)\delta_{ij} - \tau_{ij}\,, \tag{121}$$

where $[\boldsymbol{L}]$ is called the *Lighthill stress tensor*.

Indeed, we can rewrite (116) as the momentum equation for an ideal, stationary acoustic medium of mean density ρ_0 and speed of sound c_0 subjected to the externally applied stress L_{ij}

$$\frac{\partial \rho v_i}{\partial t} + \frac{\partial \pi^0_{ij}}{\partial x_j} = -\frac{\partial}{\partial x_j}\left(\pi_{ij} - \pi^0_{ij}\right)\,, \tag{122}$$

or equivalent

$$\frac{\partial \rho v_i}{\partial t} + \frac{\partial}{\partial x_j}\left(c_0^2(\rho - \rho_0)\right) = -\frac{\partial L_{ij}}{\partial x_j}\,. \tag{123}$$

By eliminating the momentum density ρv_i using (119) we arrive at **Lighthill's equation**

$$\left(\frac{1}{c_0^2}\frac{\partial^2}{\partial t^2} - \nabla\cdot\nabla\right)\left(c_0^2(\rho - \rho_0)\right) = \frac{\partial^2 L_{ij}}{\partial x_i \partial x_j}\,. \tag{124}$$

It has to be noted that $(\rho - \rho_0) = \rho'$ is a fluctuating density not being equal to the acoustic density ρ_a, but a superposition of flow and acoustic parts within flow regions.

Neglecting viscous dissipation and assuming an isentropic case, we may approximate the Lighthill tensor by

$$L_{ij} \approx \rho_0 v_i v_j \quad \text{for } \mathrm{Ma}^2 \ll 1\,. \tag{125}$$

Please note that with this assumptions, the divergence of (15) provides the following equivalence (assuming an incompressible flow $\nabla\cdot\boldsymbol{v} = 0$ and $\boldsymbol{f} = 0$)

$$\nabla\cdot\nabla p_\mathrm{ic} = -\rho_0 \frac{\partial^2 v_i v_j}{\partial x_i \partial x_j} \tag{126}$$

with the incompressible flow pressure p_ic. Therefore, we may rewrite Lighthill's inhomogeneous wave equation (124) for the fluctuating pressure p' as

$$\frac{1}{c_0^2}\frac{\partial^2 p'}{\partial t^2} - \nabla\cdot\nabla p' = \nabla\cdot\nabla p_\mathrm{ic}\,. \tag{127}$$

This equation is a quite good model for the computation of sound generated by low Mach and high Reynolds number flows.

5.2 Perturbation Equations

The acoustic/viscous splitting technique for the prediction of flow induced sound was first introduced in Hardin and Pope (1994), and afterwards many groups presented alternative and improved formulations for linear and non linear wave propagation (Shen and Sørensen 1999; Ewert and Schröder 2003; Seo and Moon 2005; Munz et al. 2007). These formulations are all based on the idea, that the flow field quantities are split into compressible and incompressible parts.

For our derivation, we introduce a generic splitting of physical quantities to the conservation equations. For this purpose, we choose a combination of the two splitting approaches introduced above and define the following

$$p = \bar{p} + p_{\mathrm{ic}} + p_{\mathrm{c}} = \bar{p} + p_{\mathrm{ic}} + p_{\mathrm{a}} \tag{128}$$

$$\boldsymbol{v} = \bar{\boldsymbol{v}} + \boldsymbol{v}_{\mathrm{ic}} + \boldsymbol{v}_{\mathrm{c}} = \bar{\boldsymbol{v}} + \boldsymbol{v}_{\mathrm{ic}} + \boldsymbol{v}_{\mathrm{a}} \tag{129}$$

$$\rho = \rho_0 + \rho_1 + \rho_{\mathrm{a}}\,. \tag{130}$$

Thereby the field variables are split into mean and fluctuating parts just like in the linearized Euler equations (LEE). In addition the fluctuating field variables are split into acoustic and non-acoustic components. Finally, the density correction ρ_1 is build in as introduced above. This choice is motivated by the following assumptions

- The acoustic field is a fluctuating field.
- The acoustic field is irrotational, i.e. $\nabla \times \boldsymbol{v}_{\mathrm{a}} = 0$.
- The acoustic field requires compressible media and an incompressible pressure fluctuation is not equivalent to an acoustic pressure fluctuation.

By doing so, we arrive for an incompressible flow at the following perturbation equations[2]

$$\frac{\partial p_{\mathrm{a}}}{\partial t} + \bar{\boldsymbol{v}} \cdot \nabla p_{\mathrm{a}} + \rho_0 c_0^2 \nabla \cdot \boldsymbol{v}_{\mathrm{a}} = -\frac{\partial p_{\mathrm{ic}}}{\partial t} - \bar{\boldsymbol{v}} \cdot \nabla p_{\mathrm{ic}} \tag{131}$$

$$\rho_0 \frac{\partial \boldsymbol{v}_{\mathrm{a}}}{\partial t} + \rho_0 \nabla\big(\bar{\boldsymbol{v}} \cdot \boldsymbol{v}_{\mathrm{a}}\big) + \nabla p_{\mathrm{a}} = 0 \tag{132}$$

with spatial constant mean density ρ_0 and speed of sound c_0. This system of partial differential equations is equivalent to the previously published ones (Ewert and Schröder 2003). The source term is the substantial derivative of the incompressible flow pressure p_{ic}. Using the acoustic scalar potential ψ_{a} and assuming a spacial constant mean density and speed of sound, we may rewrite (132) by

$$\nabla \left(\rho_0 \frac{\partial \psi_{\mathrm{a}}}{\partial t} + \rho_0\, \bar{\boldsymbol{v}} \cdot \nabla \psi_{\mathrm{a}} - p_{\mathrm{a}} \right) = 0\,, \tag{133}$$

and arrive at

$$p_{\mathrm{a}} = \rho_0 \frac{\partial \psi_{\mathrm{a}}}{\partial t} + \rho_0\, \bar{\boldsymbol{v}} \cdot \nabla \psi_{\mathrm{a}}\,. \tag{134}$$

Now, we substitute (134) into (131) and arrive at

$$\frac{1}{c_0^2} \frac{D^2 \psi_{\mathrm{a}}}{Dt^2} - \Delta \psi_{\mathrm{a}} = -\frac{1}{\rho_0 c_0^2} \frac{D p_{\mathrm{ic}}}{Dt}\,; \quad \frac{D}{Dt} = \frac{\partial}{\partial t} + \bar{\boldsymbol{v}} \cdot \nabla\,. \tag{135}$$

This convective wave equation fully describes acoustic sources generated by incompressible flow structures and its wave propagation through flowing media. In addition, instead of the original unknowns p_{a} and $\boldsymbol{v}_{\mathrm{a}}$ we have just the scalar unknown ψ_{a}. In

[2] For a detailed derivation of perturbation equations both for compressible as well as incompressible flows, we refer to Hüppe (2013).

accordance to the acoustic perturbation equations (APE), we name this resulting partial differential equation for the acoustic scalar potential as *Perturbed Convective Wave Equation* (PCWE).

Finally, it is of great interest that by neglecting the mean flow $\bar{\boldsymbol{v}}$ in (131) and (132), we arrive at the linearized conservation equations of acoustics with $\partial p_{\mathrm{ic}}/\partial t$ as a source term

$$\frac{1}{\rho_0 c_0^2}\frac{\partial p_{\mathrm{a}}}{\partial t} + \nabla \cdot \boldsymbol{v}_{\mathrm{a}} = \frac{-1}{\rho_0 c_0^2}\frac{\partial p_{\mathrm{ic}}}{\partial t} \tag{136}$$

$$\frac{\partial \boldsymbol{v}_{\mathrm{a}}}{\partial t} + \frac{1}{\rho_0}\nabla p_{\mathrm{a}} = 0\,. \tag{137}$$

As in the standard acoustic case, we apply $\partial/\partial t$ to (136) and $\nabla\cdot$ to (137) and subtract the two resulting equations to arrive at

$$\frac{1}{c_0^2}\frac{\partial^2 p_{\mathrm{a}}}{\partial t^2} - \nabla \cdot \nabla p_{\mathrm{a}} = \frac{-1}{c_0^2}\frac{\partial^2 p_{\mathrm{ic}}}{\partial t^2}\,. \tag{138}$$

We call this partial differential equation the aeroacoustic wave equation (AWE). Please note, that this equation can also be obtained by starting at Lighthill's inhomogeneous wave equation for incompressible flow, where we can substitute the second spatial derivative of Lighthill's tensor by the Laplacian of the incompressible flow pressure (see (126)) and arrive at (127). Using the decomposition of the fluctuating pressure p'

$$p' = p_{\mathrm{ic}} + p_{\mathrm{a}}\,.$$

results again into (138).

5.3 Comparison of Different Aeroacoustic Analogies

As a demonstrative example to compare the different acoustic analogies, we choose a cylinder in a cross flow, as displayed in Fig. 7. Thereby, the computational grid is just up to the height of the cylinder and together with the boundary conditions (bottom and top as well as span-wise direction symmetry boundary condition), we obtain a pseudo two-dimensional flow field. The diameter of the cylinder D is 1 m resulting with the inflow velocity of 1 m/s and chosen viscosity in a Reynolds number of 250 and Mach number of 0.2. From the flow simulations, we obtain a shedding frequency of 0.2 Hz (Strouhal number of 0.2). The acoustic mesh is chosen different from the flow mesh, and resolves the wavelength of two times the shedding frequency with 10 finite elements of second order. At the outer boundary of the acoustic domain we add a perfectly matched layer to efficiently absorb the outgoing waves. For the acoustic field computation we use the following formulations:

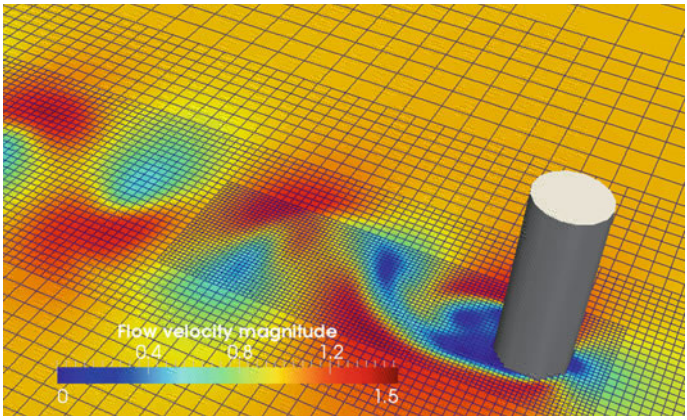


Fig. 7 Computational setup for flow computation

- Lighthill's acoustic analogy with Lighthill's tensor $[\boldsymbol{L}]$ according to (125) as source term
- Lighthill's acoustic analogy with the Laplacian of the incompressible flow pressure p_{ic} as source term (see (126))
- the aeroacoustic wave equation (AWE) according to (138)
- Perturbed Convective Wave Equation (PCWE) according to (135); for comparison, we set the mean flow velocity $\bar{\boldsymbol{v}}$ to zero.

Figure 8 displays the acoustic field for the different formulations. One can clearly see that the acoustic field of PCWE (for comparison with the other formulations we have neglected the convective terms) meets very well the expected dipole structure and is free from dynamic flow disturbances. Furthermore, the acoustic field of AWE is quite similar and exhibits almost no dynamic flow disturbances. Both computations with Lighthill's analogy show flow disturbances, whereby the formulation with the Laplacian of the incompressible flow pressure as source term shows qualitative better result as the classical formulation based on the incompressible flow velocities.

6 Vibroacoustics

In many technical applications, vibrating structures are immersed in an acoustic fluid. Therefore, acoustic waves are generated, which are acting as a surface pressure load on the vibrating structure. In general, we distinguish between the following two situations concerning mechanical-acoustic couplings:

- *Strong Coupling*: In this case, the mechanical and acoustic field equations including their couplings have to be solved simultaneously (two way coupling). A typical example is a piezoelectric ultrasound array immersed in water (see Fig. 9).

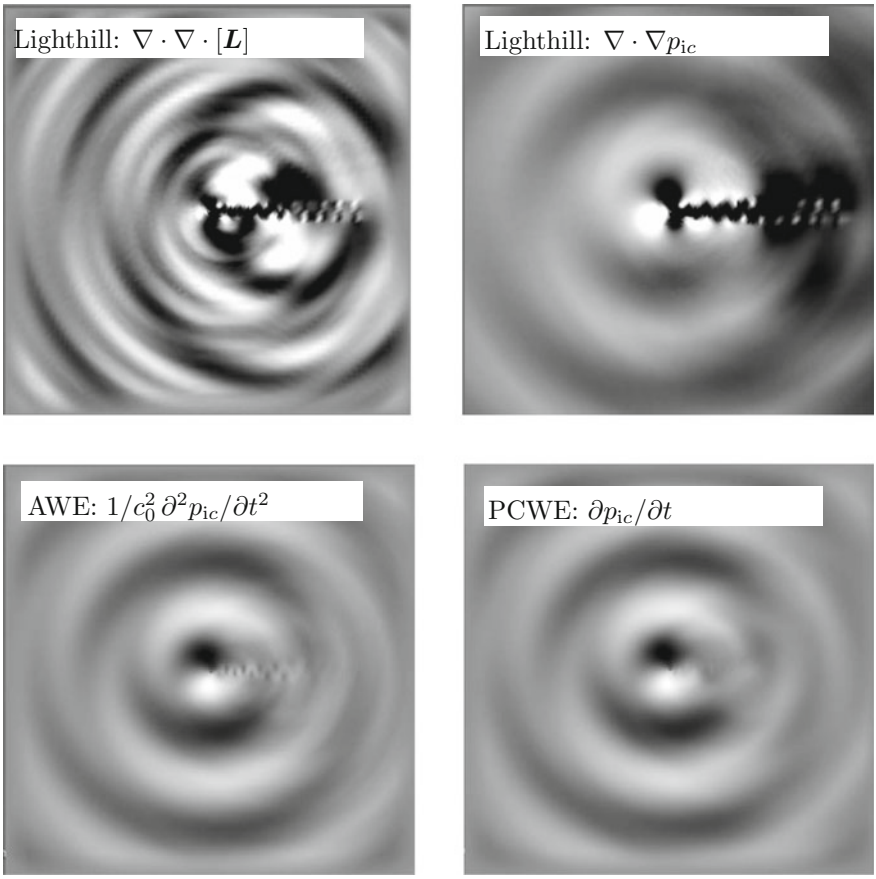


Fig. 8 Computed acoustic field with the different formulations

- *Weak Coupling*: If the pressure forces of the fluid on the solid are negligible, a sequential computation can be performed (one way coupling). For example, the acoustic sound field of an electric transformer as displayed in Fig. 10 can be obtained in this way. Thus, in a first simulation the mechanical surface vibrations are calculated, which are then used as the input for an acoustic field computation.

6.1 Interface Conditions

At a solid–fluid interface, the continuity requires that the normal component of the mechanical surface velocity of the solid must coincide with the normal component of the acoustic velocity of the inviscid fluid (see Fig. 11). Thus, the following relation

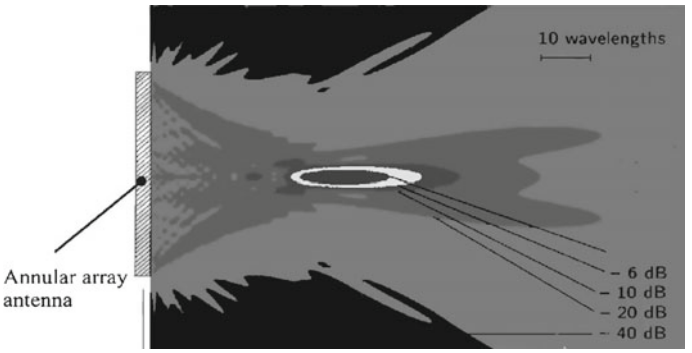


Fig. 9 Acoustic sound field of a piezoelectric ultrasound array antenna

Fig. 10 Noise radiation from the tank of an electric power transformer

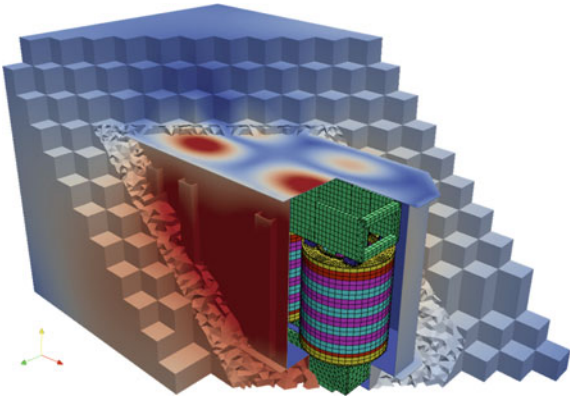
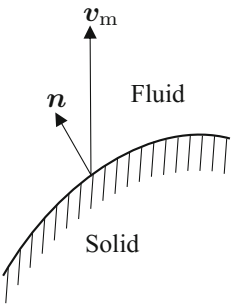


Fig. 11 Solid–fluid interface



between the velocity $\boldsymbol{v}_{\mathrm{m}}$ of the solid expressed by the mechanical displacement $\boldsymbol{u}$ and the acoustic particle velocity $\boldsymbol{v}_{\mathrm{a}}$ expressed by the acoustic scalar potential ψ_{a} arises

$$\begin{aligned} \boldsymbol{v}_{\mathrm{m}} &= \frac{\partial \boldsymbol{u}}{\partial t} \qquad \boldsymbol{v}_{\mathrm{a}} = -\nabla \psi_{\mathrm{a}} \\ \boldsymbol{n} \cdot (\boldsymbol{v}_{\mathrm{m}} - \boldsymbol{v}_{\mathrm{a}}) &= 0 \\ \boldsymbol{n} \cdot \frac{\partial \boldsymbol{u}}{\partial t} &= -\boldsymbol{n} \cdot \nabla \psi_{\mathrm{a}} = -\frac{\partial \psi_{\mathrm{a}}}{\partial \boldsymbol{n}} \,. \end{aligned} \tag{139}$$

In addition, one has to consider the fact that the ambient fluid causes on the surface a mechanical stress $\boldsymbol{\sigma}_{\mathrm{n}}$

$$\boldsymbol{\sigma}_{\mathrm{n}} = -\boldsymbol{n}\, p_{\mathrm{a}} = -\boldsymbol{n} \rho_0 \frac{\partial \psi_{\mathrm{a}}}{\partial t} \,, \tag{140}$$

which acts like a pressure load on the solid.

When modeling special wave phenomena, we often arrive at a partial differential equation for the acoustic pressure. Therewith, we will also derive the coupling conditions between the mechanical displacement and acoustic pressure at a solid–fluid interface. For the first coupling condition, the continuity of the velocities, we have to establish the relation between the acoustic particle velocity $\boldsymbol{v}_{\mathrm{a}}$ and the acoustic pressure p_{a}. According to the linearized momentum equation (see (62) and assuming zero source term), we can express the normal component of $\boldsymbol{v}_{\mathrm{a}}$ by

$$\boldsymbol{n} \cdot \frac{\partial \boldsymbol{v}_{\mathrm{a}}}{\partial t} = -\frac{1}{\rho_0} \frac{\partial p_{\mathrm{a}}}{\partial \boldsymbol{n}} \,. \tag{141}$$

Therewith, since $\boldsymbol{n} \cdot \boldsymbol{v}_{\mathrm{m}} = \boldsymbol{n} \cdot \boldsymbol{v}_{\mathrm{a}}$ holds, we get the relation to the mechanical displacement by

$$\boldsymbol{n} \cdot \frac{\partial^2 \boldsymbol{u}}{\partial t^2} = -\frac{1}{\rho_0} \frac{\partial p_{\mathrm{a}}}{\partial \boldsymbol{n}} \,. \tag{142}$$

The second coupling condition as defined in (140) is already established for an acoustic pressure formulation.

7 Appendix

Here, we provide often used operations both in vector and index notation.

- Scalar product of two vectors

$$\boldsymbol{a} \cdot \boldsymbol{b} = c \;\rightarrow\; a_i b_i = c \tag{143}$$

- Vector product of two vectors

$$\boldsymbol{a} \times \boldsymbol{b} = \boldsymbol{c} \;\rightarrow\; \epsilon_{ijk} a_j b_k = c_i \tag{144}$$

- Gradient of a scalar

$$\nabla \phi = \boldsymbol{u} \;\rightarrow\; \frac{\partial \phi}{\partial x_i} = u_i \tag{145}$$

- Gradient of a vector

$$\nabla \boldsymbol{a} = \begin{pmatrix} \frac{\partial a_1}{\partial x_1} & \frac{\partial a_2}{\partial x_1} & \frac{\partial a_3}{\partial x_1} \\ \frac{\partial a_1}{\partial x_2} & \frac{\partial a_2}{\partial x_2} & \frac{\partial a_3}{\partial x_2} \\ \frac{\partial a_1}{\partial x_3} & \frac{\partial a_2}{\partial x_3} & \frac{\partial a_3}{\partial x_3} \end{pmatrix} \rightarrow \frac{\partial a_i}{\partial x_j} \tag{146}$$

- Gradient of a second order tensor

$$\nabla [\boldsymbol{A}] = \frac{\partial [\boldsymbol{A}]}{\partial \boldsymbol{x}} = \sum_{i,j,k=1}^{3} \frac{\partial A_{ij}}{\partial x_k} \boldsymbol{e}_i \otimes \boldsymbol{e}_j \otimes \boldsymbol{e}_k \tag{147}$$

- Divergence of a vector

$$\nabla \cdot \boldsymbol{a} = b \ \rightarrow \ \frac{\partial a_i}{\partial x_i} = b \tag{148}$$

- Divergence of a second order tensor

$$\nabla \cdot [\boldsymbol{A}] = \sum_{i,j=1}^{3} \frac{\partial A_{ij}}{\partial x_j} \boldsymbol{e}_i \tag{149}$$

- Curl of a vector

$$\nabla \times \boldsymbol{a} = \boldsymbol{b} \ \rightarrow \ \epsilon_{ijk} \frac{\partial a_k}{\partial x_j} = b_i \tag{150}$$

with

$$\epsilon_{ijk} = \begin{cases} 1 & \text{if } ijk = 123,\ 231 \text{ or } 312 \\ 0 & \text{if any two indices are the same} \\ -1 & \text{if } ijk = 132,\ 213 \text{ or } 321 \end{cases}$$

- Double product or double contraction of two second order tensors

$$[\boldsymbol{A}] : [\boldsymbol{B}] = c \ \rightarrow \ A_{ij} B_{ij} = c \tag{151}$$

- Dyadic or tensor product

$$\boldsymbol{a} \otimes \boldsymbol{b} = [\boldsymbol{C}] \ \rightarrow \ a_i b_j = C_{ij} \tag{152}$$

$$[\boldsymbol{A}] \otimes \boldsymbol{b} = [\boldsymbol{C}] \ \rightarrow \ A_{ij} b_k = C_{ijk} \tag{153}$$

$$[\boldsymbol{A}] \otimes [\boldsymbol{B}] = [\boldsymbol{D}] \ \rightarrow \ A_{ij} B_{kl} = D_{ijkl} \tag{154}$$

- Trace of a tensor

$$\mathrm{tr}([\boldsymbol{A}]) = b \ \rightarrow \ A_{ii} = b\,. \tag{155}$$

Acknowledgements The author wishes to acknowledge his former Ph.D. student Andreas Hüppe for main contributions towards the derivations of the different aeroacoustic equations. Furthermore, many thanks to Stefan Schoder for proof reading and his usefull suggestions.

References

Bécache, E., Givoli, D., & Hagstrom, T. (2010). High-order absorbing boundary conditions for anisotropic and convective wave equations. *Journal of Computational Physics*, *229*(4), 1099–1129.

Berenger, J. P. (1994). A perfectly matched layer for the absorption of electromagnetic waves. *Journal of Computational Physics*, *114*, 185–200.

Dumbser, M., & Munz, C.-D. (2005). ADER discontinuous Galerkin schemes for aeroacoustics. *Comptes Rendus Mecanique*, *333*(9), 683–687.

Durst, F. (2006). *Grundlagen der Strömungsmechanik*. Berlin: Springer.

Ewert, R., & Schröder, W. (2003). Acoustic perturbation equations based on flow decomposition via source filtering. *Journal of Computational Physics*, *188*, 365–398.

Frank, H. M., & Munz, C. D. (2016). Direct aeroacoustic simulation of acoustic feedback phenomena on a side-view mirror. *Journal of Sound and Vibration*, *371*, 132–149.

Freund, J. B., Lele, S. K., & Moin, P. (2000). Direct numerical simulation of a Mach 1.92 turbulent jet and its sound field. *AIAA Journal*, *38*, 2023–2031.

Givoli, D. (2008). Computational absorbing boundaries. In S. Marburg & B. Nolte (Eds.), *Computational acoustics of noise propagation in fluids (chapter 5)* (pp. 145–166). Berlin: Springer.

Hagstrom, T., & Warburton, T. (2009). Complete radiation boundary conditions: Minimizing the long time error growth of local methods. *SIAM Journal on Numerical Analysis*, *47*(5), 3678–3704.

Hardin, J. C., & Hussaini, M. Y. (Eds.). (1992a). Computational aeroacoustics for low Mach number flows. *Computational aeroacoustics* (pp. 50–68). New York: Springer.

Hardin, J. C., & Hussaini, M. Y. (Eds.). (1992b). Regarding numerical considerations for computational aeroacoustics. *Computational aeroacoustics* (pp. 216–228). New York: Springer.

Hardin, J. C., & Pope, D. S. (1994). An acoustic/viscous splitting technique for computational aeroacoustics. *Theoretical and Computational Fluid Dynamics*, *6*, 323–340.

Howe, M. S. (1998). *Acoustics of fluid-structure interactions*. Cambridge monographs on mechanics. Cambridge: Cambridge University Press.

Hüppe, A. (2013). Spectral finite elements for acoustic field computation. Ph.D. thesis, University of Klagenfurt, Austria, 2013.

Ihlenburg, F. (1998). *Finite element analysis of acoustic scattering*. New York: Springer.

Lighthill, M. J. (1954). On sound generated aerodynamically II. Turbulence as a source of sound. *Proceedings of the Royal Society of London*, *222*, 1–A22.

Lighthill, M. J. (1952). On sound generated aerodynamically I. General theory. *Proceedings of the Royal Society of London*, *A211*, 564–587.

Munz, C. D., Dumbser, M., & Roller, S. (2007). Linearized acoustic perturbation equations for low Mach number flow with variable density and temperature. *Journal of Computational Physics*, *224*, 352–364.

Rossing, T. D. (Ed.). (2007). *Handbook of acoustics*. Berlin: Springer.

Schlichting, H., & Gersten, K. (2006). *Grenzschicht-theorie (boundary layer theory)*. Berlin: Springer.

Seo, J. H., & Moon, Y. J. (2005). Perturbed compressible equations for aeroacoustic noise prediction at low Mach numbers. *AIAA Journal*, *43*, 1716–1724.

Shen, W. Z., & Sørensen, J. N. (1999). Aeroacoustic modelling of low-speed flows. *Theoretical and Computational Fluid Dynamics*, *13*, 271–289.

Teixeira, F. L., & Chew, W. C. (2000). Complex space approach to perfectly layers: A review and some developments. *International Journal of Numerical Modelling*, *13*, 441–455.

Non-conforming Finite Elements for Flexible Discretization with Applications to Aeroacoustics

Manfred Kaltenbacher

Abstract The non-conforming Finite Element (FE) method allows the coupling of two or more sub-domains with quite different mesh sizes. Therewith, we gain the flexibility to choose for each sub-domain an optimal grid. The two proposed methods - Mortar and Nitsche-type mortaring - fulfill the physical conditions along the non-conforming interfaces. We exploit this capability and apply it to real engineering applications in aeroacoustics. The results clearly demonstrate the superiority of the non-conforming FE method over the standard FE method concerning pre-processing, mesh generation flexibility, accuracy and computational time.

1 Overview

For low Mach number flows, the speed of sound is much greater than the mean flow velocity and therefore the acoustic wavelength is much greater than the diameters of the eddies in the flow. Therefore, the only practicable approach for such cases to compute flow induced sound, known as computational aeroacoustics, is based on hybrid methods (Kaltenbacher et al. 2010). These methods compute the flow on a restricted sub-domain in a first step applying, e.g., Large Eddy Simulation (LES) to accurately resolve the main turbulent flow structures. In a subsequent step, the acoustic wave propagation within this sub-domain as well as in an ambient surrounding sub-domain is computed. The main approaches for this step are solving the linearized Euler equations, the acoustic perturbation equations, the linearized perturbed compressible equations or Lighthill's inhomogeneous wave equation (Kaltenbacher 2015). All these methods have in common, that they compute within the flow domain acoustic source terms based on the flow computation (Computational Fluid Dynamics, CFD). In order to accurately resolve these source terms, a much finer discretization is needed in the flow domain as in the ambient domain of free wave radiation.

M. Kaltenbacher (✉)
Institute of Mechanics and Mechatronics, TU Wien, Vienna, Austria
e-mail: manfred.kaltenbacher@tuwien.ac.at

M. Kaltenbacher (ed.), *Computational Acoustics*, CISM International Centre for Mechanical Sciences 579, DOI 10.1007/978-3-319-59038-7_2

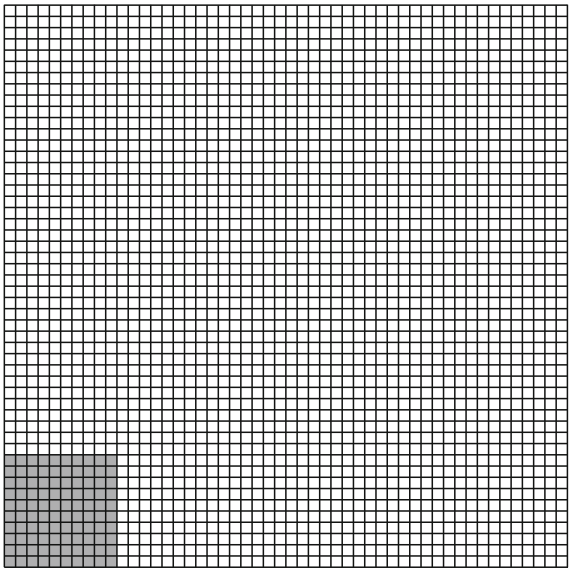


Fig. 1 Uniform mesh

The simplest approach to resolve the different grid sizes is to keep the fine discretization necessary for one sub-domain also for the other sub-domain (cf. Fig. 1). However, in many cases, a tremendous number of unknowns is obtained, so that a solution even on high performance computers is not feasible. In a second approach, the mesh could be gradually coarsened as in Fig. 2. Quite often this is the only possible choice, if the standard conforming FE method is used, since it can only handle a geometrically conforming triangulation. Unfortunately, the numerical accuracy of wave propagation applications depend very sensitively on the shape regularity of the underlying mesh. Thus, a small transition zone from fine to coarse grids results in a poor numerical approximation. Therefore, in order to meet the requirements of different mesh sizes and to gain full flexibility for the discretization, we propose to use the non-conforming FE method. More precisely the mesh-size ratio does not enter into the a priori error estimates. Using this approach, one gains much more flexibility in the modeling, since specially tuned meshes for the subproblems can be used. Most important is that the proposed formulations fulfill the physical interface conditions. Therefore, the advantages can be summarized as follows (Fig. 3):

- Pre-processing is much more flexible, since grids in the different sub-domains do not influence each other. Depending on the implementation, the global mesh may be read in parts from multiple mesh input files. This makes parameter studies handy to conduct.
- The approximation order can be chosen independently for each sub-domain. This permits to use higher order elements in regions, where the solution is known to be smooth and fine discretizations using low order elements may be used in regions, where singularities in the solution occur.

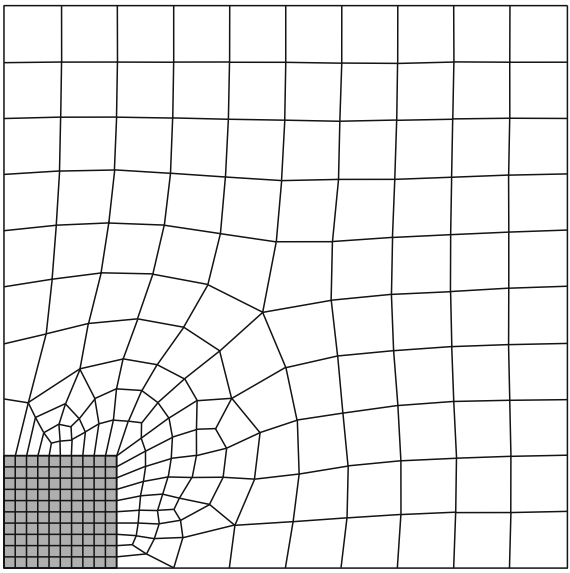


Fig. 2 Coarsening mesh

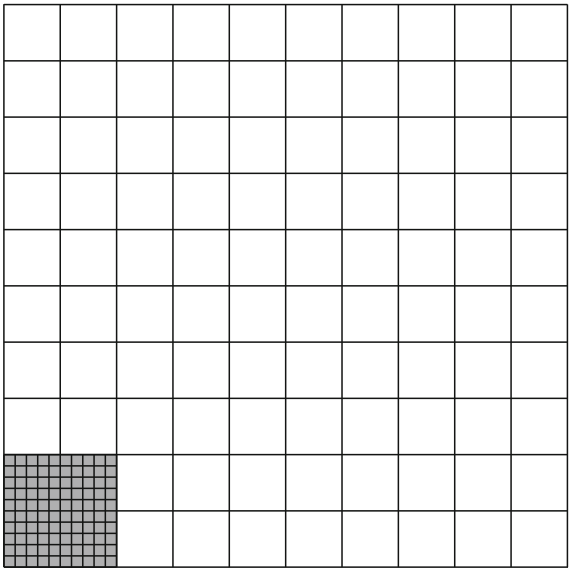


Fig. 3 Non-conforming mesh

- The method can be used for parallelization. If only a single physical field is involved, our method can be classified as a Finite Element Tearing and Interconnection dual-primal (FETI-DP) method in domain decomposition terms, see, e.g., Langer and Steinbach (2003), Dokeva (2006).

Here, we focus on computational aeroacoustics and discuss formulations and applications in case of acoustic-acoustic coupling. However, we want to note that non-conforming grid techniques are applicable to domain coupling field problems, e.g. vibro-acoustics (Flemisch et al. 2012), fluid-structure-interaction (Klöppel et al. 2011), electro-thermal coupling (Köck et al. 2015).

The non-conforming grid techniques have been implemented in our multiphysics research software CFS++ (see cfs-doc.mdmt.tuwien.ac.at), and used for the computations of the applications described in Sect. 5.

2 Non-conforming FE Formulations

We will investigate two approaches to handle non-conforming grids: (1) Mortar coupling, see, e.g., Bernardi et al. (1994), Wohlmuth (2000) and (2) Nitsche-type coupling, see, e.g., Hansbo et al. (2003), Fritz et al. (2004). In the first approach, we guarantee the strong coupling of the numerical flux (normal derivative of the acoustic pressure) by introducing a Lagrange multiplier and coupling of the acoustic pressure in a weak sense. Nitsche-type coupling does not need the additional Lagrange multiplier and handles the coupling by symmetrizing the bilinear form and adding a special jump term.

We assume a global domain Ω and its decomposition into two sub-domains Ω_1, Ω_2 as displayed in Fig. 4. Thus, in each sub-domain we solve the wave equation for the acoustic pressure $p_{ai} : \Omega_i \times (0, T) \to \mathrm{IR}$,

$$\frac{1}{c^2}\,\ddot{p}_{ai} - \Delta\, p_{ai} = g_i, \quad \text{in } \Omega_i \times (0, T),\ i = 1, 2 \tag{1}$$

completed by appropriate initial conditions at time $t = 0$ and boundary conditions on the global boundary Γ_a. In (1) a dot over a variable denotes the derivative with respect to time, i.e. $\ddot{p}_a = \partial^2 p_a / \partial t^2$. According to the physical interface conditions, we have to impose continuity for trace and flux of the acoustic pressure along the common interface Γ_I, i.e.,

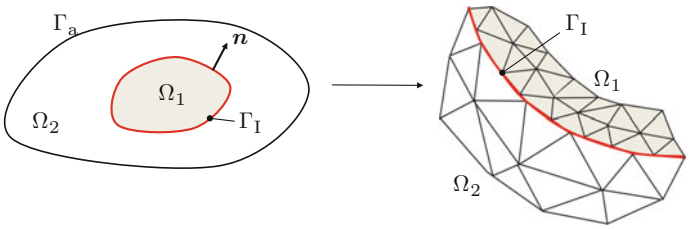


Fig. 4 Acoustic domain with two sub-domains Ω_1 and Ω_2 with different discretizations

$$p_{\mathrm{a}1} = p_{\mathrm{a}2} \;\text{ and }\; \frac{\partial p_{\mathrm{a}1}}{\partial \boldsymbol{n}} = \frac{\partial p_{\mathrm{a}2}}{\partial \boldsymbol{n}} \;\text{ on }\; \Gamma_{\mathrm{I}}\,. \tag{2}$$

Without any limitation and to keep the focus on the main steps achieving non-conforming FE formulations, we set homogeneous Dirichlet boundary condition for the acoustic pressure p_{a} at Γ_{a}.

2.1 Mortar Formulation

The flux coupling condition is enforced in a strong sense by introducing a Lagrange multiplier

$$\lambda = -\frac{\partial p_{\mathrm{a}1}}{\partial \boldsymbol{n}} = -\frac{\partial p_{\mathrm{a}2}}{\partial \boldsymbol{n}}\,. \tag{3}$$

However, the continuity of the trace will be understood in a weak sense

$$\int_{\Gamma_{\mathrm{I}}} (p_{\mathrm{a}1} - p_{\mathrm{a}2})\mu \,\mathrm{d}\boldsymbol{s} = 0 \tag{4}$$

for all test functions μ out of a suitable Lagrange multiplier space. We proceed with the weak formulation and obtain from (1)

$$\int_{\Omega_i} \frac{1}{c^2}\, w_i\, \ddot{p}_{\mathrm{a}i}\; \mathrm{d}\boldsymbol{x} + \int_{\Omega_i} \nabla w_i \cdot \nabla p_{\mathrm{a}i}\; \mathrm{d}\boldsymbol{x} - \int_{\Gamma_{\mathrm{I}}} w_i\, \boldsymbol{n}_i \cdot \nabla p_{\mathrm{a}i}\; \mathrm{d}\boldsymbol{s} = \int_{\Omega_i} w_i\, g_i\; \mathrm{d}\boldsymbol{x}\,,$$

for all test functions w_i, $i = 1, 2$. Please note that the surface term

$$\int_{\Gamma_{\mathrm{a}}} w_2\, \boldsymbol{n}_{\mathrm{a}} \cdot \nabla p_{\mathrm{a}2}\; \mathrm{d}\boldsymbol{s}$$

vanishes, since the test function is zero at Dirichlet boundaries of p_{a}. Inserting the definition of the Lagrange multiplier (3) and summing up, we obtain the symmetric evolutionary saddle point problem of finding $p_{\mathrm{a}1}$, $p_{\mathrm{a}2}$ and λ such that

$$\sum_{i=1}^{2} \left(\int_{\Omega_i} \frac{1}{c^2} w_i\, \ddot{p}_{\mathrm{a}i}\, \mathrm{d}\boldsymbol{x} + \int_{\Omega_i} \nabla w_i \cdot \nabla p_{\mathrm{a}i}\; \mathrm{d}\boldsymbol{x} \right) + \int_{\Gamma_{\mathrm{I}}} (w_1 - w_2)\lambda \,\mathrm{d}\boldsymbol{s} = \sum_{i=1}^{2} \int_{\Omega_i} w_i g_i \,\mathrm{d}\boldsymbol{x} \tag{5}$$

$$\int_{\Gamma_{\mathrm{I}}} (p_{\mathrm{a}1} - p_{\mathrm{a}2})\mu \,\mathrm{d}\boldsymbol{s} = 0 \tag{6}$$

for all μ and w_i, $i = 1, 2$. We now face a primal-dual problem, where the coupling is realized in terms of Lagrange multipliers. In a next step, we perform a spatial discretization, assume the Lagrange multiplier to be chosen with respect to Ω_1 and choose the following ansatz

$$w_1 \approx w_1^h = \sum_i N_{1i}\, w_{1i}\,; \quad p_{\mathrm{a},1} \approx p_{\mathrm{a},1}^h = \sum_j N_{1j}\, p_{\mathrm{a},1,j} \tag{7}$$

$$w_2 \approx w_2^h = \sum_i N_{2i}\, w_{2i}\,; \quad p_{\mathrm{a},2} \approx p_{\mathrm{a},2}^h = \sum_j N_{2j}\, p_{\mathrm{a},2,j} \tag{8}$$

$$\lambda \approx \lambda^h = \sum_k \phi_k\, \lambda_k\,. \tag{9}$$

Substituting this ansatz into (5), (6), results in the semi-discrete Galerkin formulation, which reads as

$$\begin{pmatrix} \boldsymbol{M}_1 & 0 & 0 \\ 0 & \boldsymbol{M}_2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \ddot{\underline{p}}_{\mathrm{a}1} \\ \ddot{\underline{p}}_{\mathrm{a}2} \\ \ddot{\underline{\lambda}} \end{pmatrix} + \begin{pmatrix} \boldsymbol{K}_1 & 0 & \boldsymbol{D} \\ 0 & \boldsymbol{K}_2 & \boldsymbol{M} \\ \boldsymbol{D}^T & \boldsymbol{M}^T & 0 \end{pmatrix} \begin{pmatrix} \underline{p}_{\mathrm{a}1} \\ \underline{p}_{\mathrm{a}2} \\ \underline{\lambda} \end{pmatrix} = \begin{pmatrix} \underline{f}_1 \\ \underline{f}_2 \end{pmatrix}. \tag{10}$$

In (10) $\boldsymbol{M}_i$ and $\boldsymbol{K}_i$ are the standard mass and stiffness matrices, see e.g. (Kaltenbacher 2015), $\underline{f}_i$ the algebraic vectors of the right hand side in Ω_i, and $\underline{p}_{\mathrm{a}1}$, $\underline{p}_{\mathrm{a}2}$, $\underline{\lambda}$ the algebraic vectors of the unknown acoustic pressures in Ω_1, Ω_2 and Lagrange multiplier along Γ_{I}, respectively. The coupling matrices $\boldsymbol{D}$, $\boldsymbol{M}$ are given by

$$\boldsymbol{D}^e = [\boldsymbol{D}_{pq}]; \quad \boldsymbol{D}_{pq} = \int_{\Gamma_{\mathrm{I}}} N_{1p}\phi_q \mathrm{d}\boldsymbol{s}, \tag{11}$$

$$\boldsymbol{M} = [\boldsymbol{M}_{pq}]; \quad \boldsymbol{M}_{pq} = \int_{\Gamma_{\mathrm{I}}} N_{2p}\phi_q \mathrm{d}\boldsymbol{s}, \tag{12}$$

where N_{1p} and N_{2p} denote the finite element basis functions on $\mathcal{T}_1$ and $\mathcal{T}_2$, respectively, and ϕ_q denotes the Lagrange multiplier basis function associated with node q. We note that the assembly of $\boldsymbol{D}$ poses no difficulty since all basis functions involved are defined with respect to the same grid $\mathcal{T}_1$. However, the assembly of $\boldsymbol{M}$ is more involved, since N_{2p} and ϕ_q are defined with respect to different grids (see Sect. 4).

2.2 Nitsche-Type Mortaring Formulation

To handle the non-conforming discretization within Nitsche's method, we start at the weak formulation for both sub-domains Ω_1 and Ω_2

$$\int_{\Omega_1} \frac{1}{c^2}\, w_1 \ddot{p}_{a1} \mathrm{d}\boldsymbol{x} + \int_{\Omega_1} \nabla w_1 \cdot \nabla p_{a1} \mathrm{d}\boldsymbol{x} - \int_{\Gamma_I} w_1 \frac{\partial p_{a1}}{\partial \boldsymbol{n}_1}\, \mathrm{d}\boldsymbol{s} = \int_{\Omega_1} w_1 g_1\, \mathrm{d}\boldsymbol{x} \qquad (13)$$

$$\int_{\Omega_2} \frac{1}{c^2}\, w_2 \ddot{p}_{a2} \mathrm{d}\boldsymbol{x} + \int_{\Omega_2} \nabla w_2 \cdot \nabla p_{a2} \mathrm{d}\boldsymbol{x} - \int_{\Gamma_I} w_2 \frac{\partial p_{a2}}{\partial \boldsymbol{n}_2}\, \mathrm{d}\boldsymbol{s} = \int_{\Omega_2} w_2 g_2\, \mathrm{d}\boldsymbol{x}\,. \qquad (14)$$

In a next step, we add the two Eqs. (13) and (14), and explore the relation

$$\boldsymbol{n} = \boldsymbol{n}_1 = -\boldsymbol{n}_2\,; \quad \frac{\partial p_{a1}}{\partial \boldsymbol{n}_1} = \frac{\partial p_{a1}}{\partial \boldsymbol{n}} = \frac{\partial p_{a2}}{\partial \boldsymbol{n}_2} = -\frac{\partial p_{a2}}{\partial \boldsymbol{n}}$$

to arrive at

$$\int_{\Omega_1} \frac{1}{c^2}\, w_1 \ddot{p}_{a1} \mathrm{d}\boldsymbol{x} + \int_{\Omega_1} \nabla w_1 \cdot \nabla p_{a1} \mathrm{d}\boldsymbol{x} + \int_{\Omega_2} \frac{1}{c^2}\, w_2 \ddot{p}_{a2} \mathrm{d}\boldsymbol{x} + \int_{\Omega_2} \nabla w_2 \cdot \nabla p_{a2} \mathrm{d}\boldsymbol{x} - \int_{\Gamma_I} [w] \frac{\partial p_{a1}}{\partial n}\, \mathrm{d}\boldsymbol{s} = \int_{\Omega_1} w_1 g_1\, \mathrm{d}\boldsymbol{x} + \int_{\Omega_2} w_2 g_2\, \mathrm{d}\boldsymbol{x}\,. \qquad (15)$$

In (15) the operator [ ] defines the jump operator, e.g., $[w] = w_1 - w_2$. In order to retain symmetry, we add to (15) the following term

$$-\int_{\Gamma_I} [p_a] \frac{\partial w_1}{\partial n}\, \mathrm{d}\boldsymbol{s} \quad \text{with} \quad [p_a] = p_{a1} - p_{a2}\,.$$

This operation is allowed, since $[p_a]$ is forced to be zero at the interface. In a final step, we add along the interface Γ_I the term

$$\beta \sum_E \frac{1}{h_E} \int_{\Gamma_E} [p_a]\,[w]\, \mathrm{d}\boldsymbol{s} \qquad (16)$$

with β the penalty factor. In (16) h_E is a characteristic length scale of each interface element E (space discrete level). Therewith, we arrive at the following final formulation for Nitsche-type mortaring

$$\int_{\Omega_1} \frac{1}{c^2}\, w_1 \ddot{p}_{a1} \mathrm{d}\boldsymbol{x} + \int_{\Omega_1} \nabla w_1 \cdot \nabla p_{a1} \mathrm{d}\boldsymbol{x} + \int_{\Omega_2} \frac{1}{c^2}\, w_2 \ddot{p}_{a2} \mathrm{d}\boldsymbol{x} + \int_{\Omega_2} \nabla w_2 \cdot \nabla p_{a2} \mathrm{d}\boldsymbol{x} \underbrace{- \int_{\Gamma_I} [w] \frac{\partial p_{a1}}{\partial n}\, \mathrm{d}\boldsymbol{s}}_{Consistency} \underbrace{- \int_{\Gamma_I} [p_a] \frac{\partial w_1}{\partial n}\, \mathrm{d}\boldsymbol{s}}_{Symmetrization}$$

$$\underbrace{+\beta \sum_E \frac{1}{h_E} \int_{\Gamma_E} [p_{\mathrm{a}}]\,[w]\,\mathrm{d}\boldsymbol{s}}_{Penalty/Stabilization} = \int_{\Omega_1} w_1 g_1\,\mathrm{d}\boldsymbol{x} + \int_{\Omega_2} w_2 g_2\,\mathrm{d}\boldsymbol{x}\,. \tag{17}$$

If the penalty parameter β is chosen large enough, the bilinear form is coercive on the discrete space and one derives optimal a priori error estimates in both the energy norm and the L_2 norm for polynomials of arbitrary degree (Hansbo et al. 2003). In a next step, we perform a spatial discretization according to (7), (8) and arrive at

$$\begin{aligned}
&\int_{\Omega_1} \frac{1}{c^2}\, w_1^h\, \ddot{p}_{\mathrm{a}1}^h\,\mathrm{d}\boldsymbol{x} + \int_{\Omega_1} \nabla w_1^h \cdot \nabla p_{\mathrm{a}1}^h\,\mathrm{d}\boldsymbol{x} - \int_{\Gamma_{\mathrm{I}}} w_1^h \frac{\partial p_{\mathrm{a}1}^h}{\partial \boldsymbol{n}}\,\mathrm{d}\boldsymbol{s} \\
&- \int_{\Gamma_{\mathrm{I}}} \frac{\partial w_1^h}{\partial \boldsymbol{n}}\, p_{\mathrm{a}1}^h\,\mathrm{d}\boldsymbol{s} + \int_{\Gamma_{\mathrm{I}}} \frac{\partial w_1^h}{\partial \boldsymbol{n}}\, p_{\mathrm{a}2}^h\,\mathrm{d}\boldsymbol{s} + \beta \sum_{E(\Gamma_{\mathrm{I}})} \frac{1}{h_E} \int_{\Gamma_{\mathrm{I}}} w_1^h\, p_{\mathrm{a}1}^h\,\mathrm{d}\boldsymbol{s} \\
&\qquad - \beta \sum_{E(\Gamma_{\mathrm{I}})} \frac{1}{h_E} \int_{\Gamma_{\mathrm{I}}} w_1^h\, p_{\mathrm{a}2}^h\,\mathrm{d}\boldsymbol{s} = \int_{\Omega_1} w_1^h g_1\,\mathrm{d}\boldsymbol{x}
\end{aligned} \tag{18}$$

$$\begin{aligned}
&\int_{\Omega_2} \frac{1}{c^2}\, w_2^h\, \ddot{p}_{\mathrm{a}2}^h\,\mathrm{d}\boldsymbol{x} + \int_{\Omega_2} \nabla w_2^h \cdot \nabla p_{\mathrm{a}2}^h\,\mathrm{d}\boldsymbol{x} + \int_{\Gamma_{\mathrm{I}}} w_2^h \frac{\partial p_{\mathrm{a}1}^h}{\partial \boldsymbol{n}}\,\mathrm{d}\boldsymbol{s} \\
&+ \beta \sum_{E(\Gamma_{\mathrm{I}})} \frac{1}{h_E} \int_{\Gamma_{\mathrm{I}}} w_2^h\, p_{\mathrm{a}2}^h\,\mathrm{d}\boldsymbol{s} - \beta \sum_{E(\Gamma_{\mathrm{I}})} \frac{1}{h_E} \int_{\Gamma_{\mathrm{I}}} w_2^h\, p_{\mathrm{a}1}^h\,\mathrm{d}\boldsymbol{s} \\
&\qquad = \int_{\Omega_2} w_2^h g_2\,\mathrm{d}\boldsymbol{x}\,.
\end{aligned} \tag{19}$$

In matrix notation, the discrete system of equations reads as

$$\begin{aligned}
&\begin{pmatrix} \boldsymbol{M}_1 & 0 \\ 0 & \boldsymbol{M}_2 \end{pmatrix} \begin{pmatrix} \ddot{\underline{p}}_{\mathrm{a}1} \\ \ddot{\underline{p}}_{\mathrm{a}2} \end{pmatrix} + \begin{pmatrix} \boldsymbol{K}_1 & 0 \\ 0 & \boldsymbol{K}_2 \end{pmatrix} \begin{pmatrix} \underline{p}_{\mathrm{a}1} \\ \underline{p}_{\mathrm{a}2} \end{pmatrix} \\
&\qquad + \begin{pmatrix} \boldsymbol{K}_{\Gamma_{\mathrm{I}1}} & \boldsymbol{K}_{\Gamma_{\mathrm{I}1}\Gamma_{\mathrm{I}2}} \\ \boldsymbol{K}_{\Gamma_{\mathrm{I}2}\Gamma_{\mathrm{I}1}} & \boldsymbol{K}_{\Gamma_{\mathrm{I}2}} \end{pmatrix} \begin{pmatrix} \underline{p}_{\mathrm{a}1} \\ \underline{p}_{\mathrm{a}2} \end{pmatrix} = \begin{pmatrix} \underline{f}_1 \\ \underline{f}_2 \end{pmatrix}.
\end{aligned} \tag{20}$$

Thereby, $\boldsymbol{M}_k$ and $\boldsymbol{K}_k$ are the standard mass and stiffness matrices, respectively. The additional matrices according to the interface compute as follows

$$\begin{aligned}
\boldsymbol{K}_{\Gamma_{\mathrm{I}1}}^{ij} &= -\int_{\Gamma_{\mathrm{I}1}} N_{1i} \frac{\partial N_{1j}}{\partial \boldsymbol{n}}\,\mathrm{d}\boldsymbol{s} - \int_{\Gamma_{\mathrm{I}1}} \frac{\partial N_{1i}}{\partial \boldsymbol{n}}\, N_{1j}\,\mathrm{d}\boldsymbol{s} \\
&\quad + \beta \sum_{E(\Gamma_{\mathrm{I}1})} \frac{1}{h_E} \int_{\Gamma_E} N_{1i} N_{1j}\,\mathrm{d}\boldsymbol{s}
\end{aligned} \tag{21}$$

$$\boldsymbol{K}^{ij}_{\Gamma_{I1}\Gamma_{I2}} = \int_{\Gamma_{I1}} \frac{\partial N_{1i}}{\partial \boldsymbol{n}} N_{2j}\, \mathrm{d}\boldsymbol{s} - \beta \sum_{E(\Gamma_{I1})} \frac{1}{h_E} \int_{\Gamma_E} N_{1i} N_{2j}\, \mathrm{d}\boldsymbol{s}$$

$$= \left(\boldsymbol{K}^{ij}_{\Gamma_{I2}\Gamma_{I1}}\right)^t \tag{22}$$

$$\boldsymbol{K}^{ij}_{\Gamma_{I2}} = \beta \sum_{E(\Gamma_{I2})} \frac{1}{h_E} \int_{\Gamma_E} N_{2i} N_{2j}\, \mathrm{d}\boldsymbol{s} \tag{23}$$

$$\underline{f}^i_1 = \int_{\Omega_1} N_{1i}\, g_1\, \mathrm{d}\boldsymbol{x}\,; \quad \underline{f}^i_2 = \int_{\Omega_2} N_{2i}\, g_2\, \mathrm{d}\boldsymbol{x}\,. \tag{24}$$

Here, we have already substituted Γ_I by Γ_{I1} as well as Γ_{I2}, which are the discretized interfaces of Ω_1 and Ω_2. Furthermore, the computation of the matrices in (21) and (22) involves basis functions N_1 and N_2, which are defined on different grids. Therefore, grid intersection operations as for the classical Mortar formulation are necessary, see Sect. 4. In addition, we note that Nitsche-type mortaring is equivalent to an IP-DG (Internal Penalty - Discontinuous Galerkin) ansatz along the non-conforming interface Γ_I. Finally, we want to emphasis that both approaches, classical Mortar and Nitsche-type mortaring, are powerful methods to correctly handle non-conforming grids both from a physical and mathematical point of view.

3 Time Discretization

In a final step to arrive at the full discrete system of equations, we have to perform a time discretization. Spurious waves, which are not resolved by the discretization (both in space and time), deteriorate the numerical solution and should be numerically damped. Since numerical damping cannot be introduced in the classical Newmark method without degrading the order of accuracy, we advise to apply a time-stepping scheme with controlled numerical dispersion such as the HHT (Hilber–Hughes–Taylor) method. Thereby, three parameters define the behavior of the time-stepping scheme: α_{HHT}, β_{HHT} and γ_{HHT}. Figure 5 demonstrates the damping behavior of different schemes. As can be seen, the standard trapezoidal scheme introduces no numerical damping. The Newmark scheme, which is just second order accurate for the parameters $\beta_{\mathrm{NM}} = 0.25$ and $\gamma_{\mathrm{NM}} = 0.5$, is able to achieve appropriate numerical damping by degrading to first order accuracy. The HHT method is unconditional stable and 2nd order accurate for $\alpha_{\mathrm{HHT}} \in [-0.3,\ 0]$ and according to the choice of this parameter introduces numerical damping. The two other parameters compute as

$$\beta_{\mathrm{HHT}} = \frac{(1-\alpha_{\mathrm{HHT}})^2}{4}\,; \quad \gamma_{\mathrm{HHT}} = \frac{(1-2\alpha_{\mathrm{HHT}})}{2}\,.$$

For a detailed analysis, we refer to Hughes (2000).

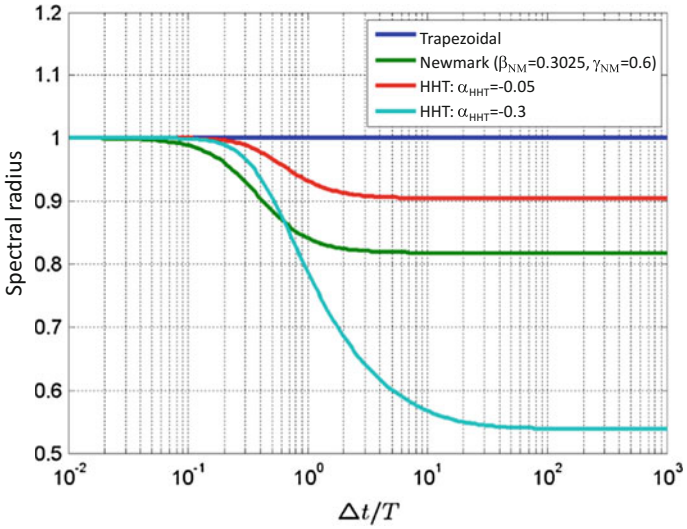


Fig. 5 Spectral radius (defined by the largest eigenvalue of the amplification matrix) over the ratio of time step size Δt to time period T

4 Mesh Intersection Operations

The feature which makes the Mortar and Nitsche-type mortaring so flexible, namely the usage of non-conforming meshes in different sub-domains, comes at the cost of a more elaborate implementation. Since the grids are allowed to be non-conforming on the interfaces of two sub-domains, the integrals defined on these interfaces involving basis functions from both sides have to be evaluated with respect to two different meshes. The decomposition of the global domain is done in a geometrically conforming way however. This guarantees that any interface inherits the discretizations of its neighboring sub-domains. It is necessary to compute the domains, where pairs of elements on the interface intersect. The corresponding integrals are then evaluated over these domains of intersection and it is up to the assembly operator to assemble the corresponding results into the correct positions of the coupling matrices.

In the following we denote the interface between two sub-domains Ω_j and Ω_k by Γ_{jk}. The triangulations corresponding to Ω_j and Ω_k are labeled $\mathcal{T}_j$ and $\mathcal{T}_k$. The nodal basis functions on $\mathcal{T}_j$ shall be denoted by N_{ja} and the ones defined on $\mathcal{T}_k$ are N_{kb}. An integral over the interface may then be written in terms of the basis function as

$$\int_{\Gamma_{jk}} N_{ja}\, N_{kb}\, \mathrm{d}\boldsymbol{s}. \tag{25}$$

For numerically evaluating this integral, we first have to determine the subsets of the interface, where pairs of elements intersect. In 2D the interfaces between sub-

domains are curves. Therefore, we have to consider the intersection of line elements in this case. If the interface is planar these are simple interval checks. If the interface is curved, we first have to project the elements onto a common line segment and do the interval checks there. These considerations also apply in a modified way for domains in 3D, where interfaces are surfaces. We have to note however that the seemingly simple operation of finding the intersection domain of arbitrary surface elements is a highly non-trivial task even for first order elements with straight edges. The last named case is however closely related to a problem in computer graphics. There 2D polygons generated during the rendering of 3D scenes have to be clipped against a view-port (cf. Greiner and Hormann 1998). Strategies and algorithms for dealing with the mesh intersection problem have been sought after for a long time in the area of domain decomposition. A small selection of available methods can be found in Puso and Laursen (2002), Park and Felippa (2002), Heinstein and Laursen (2003), Puso (2004).

If no neighborhood information between the elements on the interfaces is available and if a naive approach is taken, the operation of finding the intersection domains of all pairs of elements is of complexity $O(m \cdot n)$. Here m is the number of elements on the master side and n is the number of elements on the slave side. The operation is so expensive, since every element on the slave side has to be checked for intersection with every element on the master side and no assumptions about neighborhood are being made.

By applying space partitioning algorithms the required effort for this operation may however be drastically reduced. If neighborhood information is present in addition, an advancing front algorithm may be applied which is described in Gander and Japhet (2009). It starts at a known pair of intersecting elements and then proceeds with the intersection checks at the neighboring elements. The algorithm therewith achieves linear complexity. This improvement is of crucial importance when dealing with applications, like rotating domains (Kaltenbacher et al. 2016a), where the intersection domains have to be recomputed after each time step.

The final step is to compute the value of the integrals on the intersection domains. The method we describe here has shown to be robust and can be implemented also in FE codes, which do not provide analytical parameterizations of the domain geometry.

4.1 Intersection of Two Line Elements

If an intersection of two co-linear line elements exists, it is again a line element sharing two of the four endpoints of both parent elements in the co-linear case. To check for an intersection one has to project the endpoints $[\mathbf{m}_1, \mathbf{m}_2]$ in two dimensional coordinates of the element on the master side of the interface to the one dimensional local coordinate system defined by the endpoints of the slave element $[\mathbf{s}_1, \mathbf{s}_2]$.

The local coordinates of the slave nodes $[\mathbf{s}_1, \mathbf{s}_2]$ are trivially given by 0 and 1. The four local coordinates of the pair of lines are then brought into ascending order

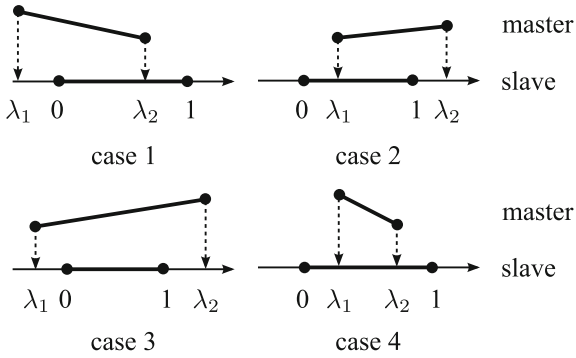


Fig. 6 Four possible cases of two lines intersecting each other

and therefore four possible cases for the intersection of two line elements may be identified as depicted in Fig. 6:

1. $\lambda_1 \leq 1 \ \wedge \ 0 < \lambda_2 < 1$: the intersection is the line $[\mathbf{s}_1, \mathbf{m}_2]$
2. $0 < \lambda_1 < 1 \ \wedge \ \lambda_2 \geq 1$: the intersection is the line $[\mathbf{m}_1, \mathbf{s}_2]$
3. $\lambda_1 \leq 0 \ \wedge \ \lambda_2 \geq 1$: the intersection is the line $[\mathbf{s}_1, \mathbf{s}_2]$
4. $\lambda_1 > 0 \ \wedge \ \lambda_2 < 1$: the intersection is the line $[\mathbf{m}_1, \mathbf{m}_2]$

We note that new points have to be generated at the projection positions on the slave element for curved interfaces in the cases 1, 2 and 4.

4.2 Intersection of Two Axis-Parallel Quadrilateral Elements

The algorithm for finding intersections of lines can be extended in a straight forward manner to a 3D setting if only axis-parallel quadrilateral elements are present on the interface. The term axis-parallel does not refer to the global coordinate axes in this context, but to the fact that the quadrilateral edges on both sides of the interface have to be parallel. This includes the case of parallelogram-shaped quadrilaterals as depicted in Fig. 7. We again compute the local coordinates (λ_1, μ_1) and (λ_2, μ_2) of the first and third corner of the master element in respect to the slave element. After bringing the local coordinates for both directions into ascending order there are sixteen possible cases for the intersection of two quadrilaterals. The ordering is necessary due to the fact that the order of nodes for elements is just guaranteed to be counter-clockwise, but the master element might have been rotated in respect to the slave element as a whole. In addition, we have to mention however that there exist many more possibilities of intersection for pairs of triangles, pairs of triangles and quadrilaterals or pairs of arbitrary shaped quadrilaterals than for the simple configuration given here. These situations require a more sophisticated treatment. A description of the algorithm used to treat arbitrary element types on curved interfaces can be found in Grabinger (2007).

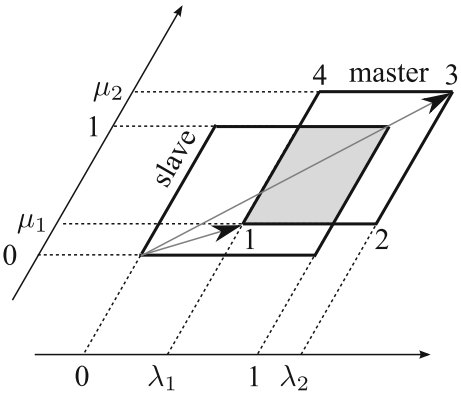


Fig. 7 Intersection of two axis-parallel parallelogram-shaped elements

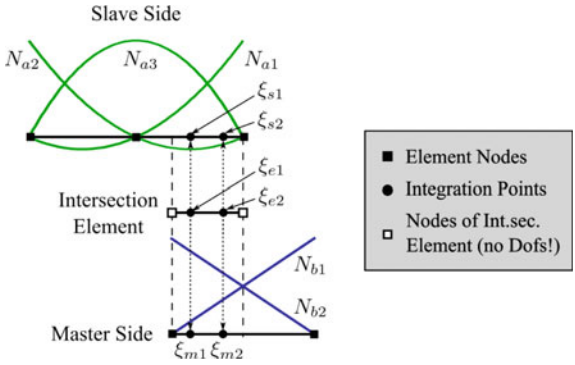


Fig. 8 Projection of quadrature points from the intersection element into the master element which is of first order and into the slave element which is of second order in this example

4.3 Evaluation of the Coupling Integrals

Once the intersection elements have been found, the coupling integral (25) can be evaluated on these elements by means of standard Gauss quadrature

$$\int_{\Gamma_{jk}} N_{ja}\, N_{kb}\, \mathrm{d}\boldsymbol{s} = \sum_{e=1}^{n_{\text{isec}}} \int_{\Gamma_e} N_a\, N_b\, \mathrm{d}\boldsymbol{s} \approx$$

$$\approx \sum_{e=1}^{n_{\text{isec}}} \sum_{l=1}^{n_{\text{int}}} W_l N_a(\boldsymbol{\xi}_l^m) N_b(\boldsymbol{\xi}_l^s) \mathcal{J}^e(\boldsymbol{\xi}_l^e)\ . \tag{26}$$

Here n_{isec} is the number of intersection elements, n_{int} is the number of quadrature points, W_l are the quadrature weights and the determinant of the Jacobian $\mathcal{J}^e$ accounts for the element mapping. The difficulty which arises when this quadrature formula is applied, is that only the quadrature point $\boldsymbol{\xi}_l^e$ in respect to the local coordinates of the intersection element is known in advance and that the points $\boldsymbol{\xi}_l^m$ in the master element and $\boldsymbol{\xi}_l^s$ in the slave element have to be projected into those elements, before the basis functions can be evaluated there (see Fig. 8). It is very important to notice that nodes

of the intersection element do not carry any degrees of freedom by themselves. The intersection element is just an auxiliary geometrical entity, which only serves as integration domain. The projection operation for general elements involves the following steps:

1. Map local coordinates $\boldsymbol{\xi}_l^e$ of quadrature point in intersection element to global coordinates
2. Map global coordinates of quadrature point to local coordinates $\boldsymbol{\xi}_l^m$ of master element
3. Map global coordinates of quadrature point to local coordinates $\boldsymbol{\xi}_l^s$ of slave element

Points 2 and 3 in general involve the application of a Newton–Raphson algorithm. A linear mapping algorithm may only be used for 2-node isoparametric line elements, 3-node isoparametric triangle elements or higher order elements which just use a linear local-to-global mapping. Once the values of the basis functions N_a and N_b have been obtained and (26) has been evaluated, the assembly operator has to make sure, that the contribution gets added to the corresponding entry in the coupling matrix.

5 Application to Aeroacoustics

5.1 Cylinder in Cross Flow

As a first practical application, the generation of sound due to a cylinder mounted on a plate in cross-flow is investigated. This simple geometry is nonetheless interesting to analyze since variations of it are very common sources for flow induced noise (e.g. antennas on cars, flagpoles, etc.). Understanding the mechanisms of sound generation for this geometry may therefore give important hints to engineers on how to reduce noise levels for similar settings. The described setting has already been subject to closer empirical and numerical investigations (c.f. Escobar 2007; Hahn 2008).

For the cylinder with rectangular cross-section having a side-length D of 20 mm (see Fig. 9), the first occurring main frequency is in the range from 50 to 60 Hz at a mean flow velocity of 10 m/s. Given the speed of sound in air at standard conditions ($c = 343$ m/s) results in a wavelength λ of about 5.72 m. Resolving the wavelength by 20 finite elements with basis functions of 1st order results in an edge length of the finite elements, which corresponds to the dimensions of the domain, on which the flow is computed. This fact alone motivates the usage of non-conforming grids at the interface towards the acoustic propagation domain.

CFD The domain, on which the flow is computed and which corresponds to the acoustic source domain, is displayed in Fig. 9. Thereby, the research program FASTEST-3D (Durst and Schäfer 1996) and the commercial software ANSYS-CFX are applied for the flow computation. The boundary conditions used in the fluid

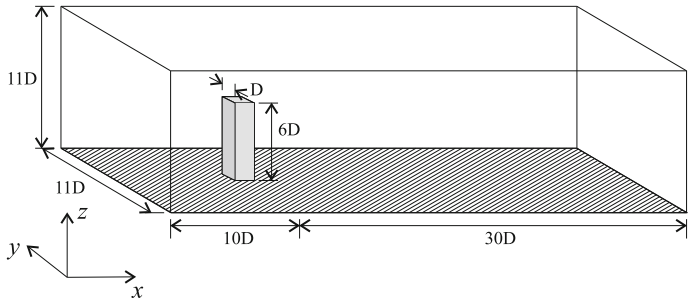


Fig. 9 Numerical domain used for fluid computations depicting dimensions. D = 20 mm

Table 1 Boundary conditions used for fluid computations

Position	Boundary condition
X = 0	Inlet profile based on experiments
X = 40D	Convective exit boundary
Z = 11D	Symmetry boundary condition
Y = 0, Y = 11D	Symmetry boundary condition
Cylinder surface and bottom	No slip boundary condition

computation with respect to the configuration from Fig. 9 are described in Table 1. Therewith, we have used a measured inflow profile with a mean velocity of 10 m/s resulting in a Reynolds number of about 13.000. FASTEST-3D uses a LES (Large Eddy Simulation) turbulence model to accurately resolve the flow structure. After a grid study, the computations have been performed on a grid with 3.1 million cells having strong refinements at the critical regions close to the cylinder and the wall. The nearest grid point in dimensionless wall coordinates is at $y^+ = 2$. The time step size was set to $\Delta t_\mathrm{f}^{LES} = 10\,\mu\mathrm{s}$, which guaranteed a resolution of up to 10 kHz, and which resulted in a CFL-number of 2.1.

For the simulation of the flow using the code ANSYS-CFX, a turbulence modeling approach based on SAS (Scale Adaptive Simulation) was employed (Menter and Egorov 2005). The SAS approach allowed us to coarsen the grid of the LES computations to about 1.1 million, which resulted in a shorter computational time and less memory usage. Regarding the time discretization, a time step size of $\Delta t_\mathrm{f}^{SAS} = 2\Delta t_\mathrm{f}^{LES} = 20\,\mu\mathrm{s}$ was used.

To get an impression about the flow field, we show in Fig. 10 the flow structure as obtained by ANSYS-CFX for a characteristic time step. The displayed results are iso-surfaces of $\boldsymbol{\omega}^2 - \boldsymbol{\epsilon}^2 = 100{,}000\,\mathrm{s}^{-2}$ colored with the eddy viscosity (here $\boldsymbol{\omega}$ representing the vorticity and $\boldsymbol{\epsilon}$ the strain rate). One can clearly see the horseshoe, the roof and span-wise vortex structure. In studying animations of the flow structure, one can observe a strong interaction between the roof and span-wise vortex, which results in a reduced vortex street behind the cylinder. For a quantitative comparison

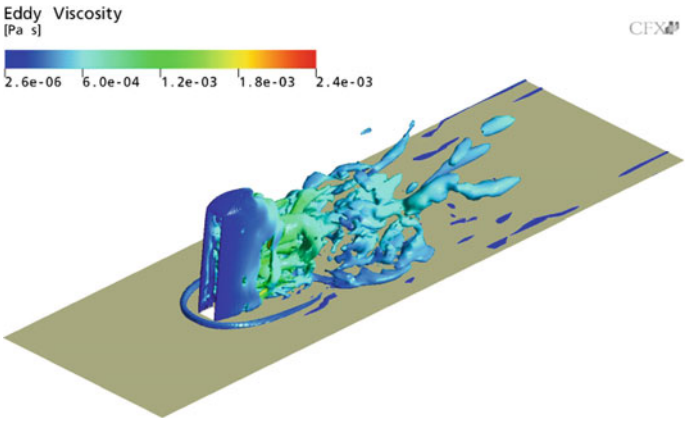


Fig. 10 Instantaneous visualization of transient flow field using SAS turbulence modeling

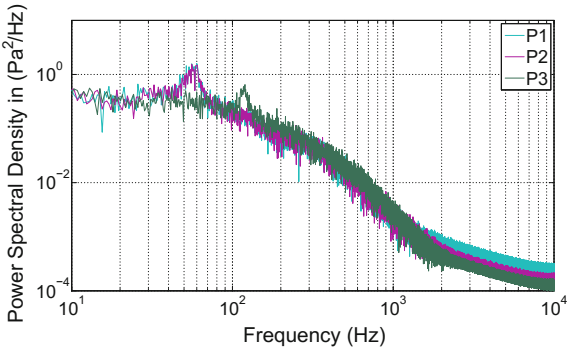


Fig. 11 Frequency spectra of the wall pressure fluctuation at different monitor points obtained by LES

between LES and SAS computations we show in Figs. 11 and 12 the spectra of the wall pressure fluctuations at different monitor points as listed in Table 2. In both simulations, pressure fluctuations on the side walls (monitoring points P1 and P2) of the cylinder show the characteristic vortex shedding frequency of about 55 Hz, which are in good agreement with experiments (Becker et al. 2008). In addition, the pressure fluctuations at monitoring point P3, which is located on the bottom behind the cylinder, exhibit in both simulations a dominant frequency at twice the vortex shedding frequency. At this point it should be noted that for both LES- and SAS-based data no significant differences were found in the acoustic field.

Acoustics The computational domain for acoustics, as it is depicted in Fig. 13, consists of the source domain, a propagation domain and a Perfectly Matched Layer (PML) to account for free field radiation (Kaltenbacher 2015). On the bottom plane as well as on the faces of the cylinder sound-hard walls are modeled by applying homogeneous Neumann boundary conditions. Here, we solve the inhomogeneous wave equation of Lighthill in the frequency domain

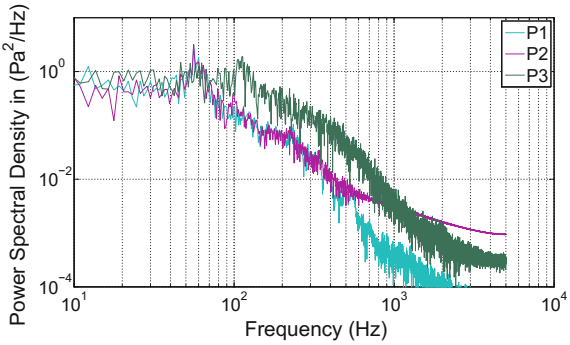


Fig. 12 Frequency spectra of the wall pressure fluctuation at different monitor points obtained by SAS

Table 2 Points at which we have evaluated the wall pressure spectra (see Fig. 9)

Position	P01	P02	P03
X	10.5 D	10.5 D	15.0 D
Y	5.0 D	6.0 D	5.5 D
Z	3.0 D	3.0 D	0.0 D

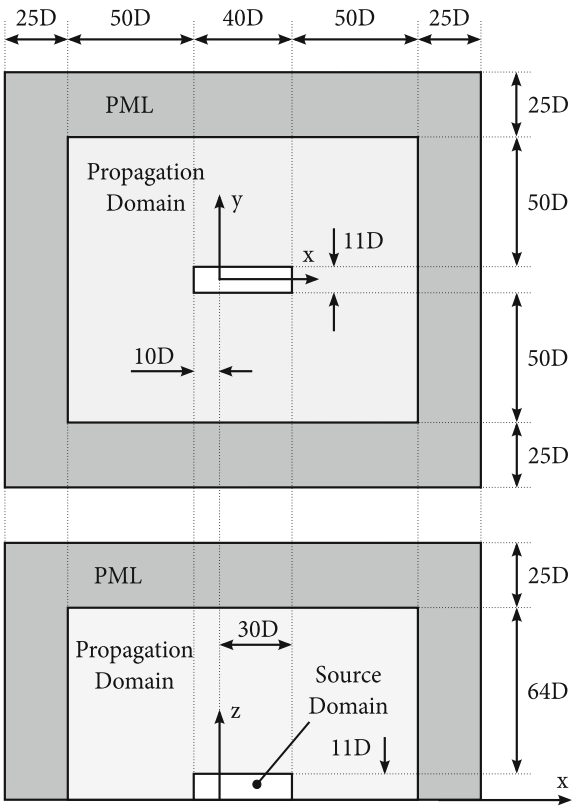


Fig. 13 Geometry of acoustic domain for harmonic simulation

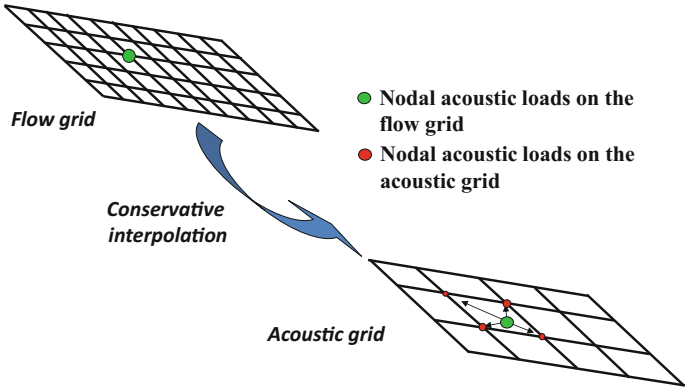


Fig. 14 Conservative interpolation from a fine CFD grid to a coarser acoustic grid

$$\frac{\partial^2 \hat{p}'}{\partial x_i^2} + k^2 \hat{p}' = -\frac{\partial^2 \hat{T}_{ij}}{\partial x_i \partial x_j} \tag{27}$$

with the harmonic pressure fluctuation $\hat{p}'$, the wavenumber k and the Fourier transformed entries $\hat{T}_{ij}$ of Lighthill's tensor. Due to the low Mach number, we approximate the entries $T_{ij}(t)$ by

$$T_{ij}(t) \approx \rho_0 v_i(t) v_j(t) \tag{28}$$

with the mean density ρ_0 and flow velocity $\boldsymbol{v}$. Here, we apply the proposed Mortar formulation, which allows to combine different meshes for the source and propagation domains and flexibly build up a global mesh specially suited for the aeolian tones which are expected in the analysis.

A crucial point for each hybrid aeroacoustic approach is the transformation of the acoustic sources from the flow grid to the acoustic grid. In order to preserve the acoustic energy, we perform an integration over the source volume (corresponding to the computational flow region) within the FE formulation and project the results to the nodes of the fine flow grid, which has to be interpolated to the coarser acoustic grid (see Fig. 14). Thereby, our interpolation has to be conservative in order to preserve the total acoustic energy. As illustrated in Fig. 14, we have to find for each nodal source F_k^{f} of the flow grid in which finite element of the acoustic grid it is located. Then, we compute from the global position $\boldsymbol{x}_k$ its local position $\boldsymbol{\xi}_k$ in the reference element. This is in the general case a nonlinear mapping and is solved by a Newton scheme. Now, with these data we can perform a bilinear interpolation and add the contribution of F_k^{f} to the nodes of the acoustic grid by using the standard finite element basis functions N_i (Kaltenbacher et al. 2010)

$$F_i^{\mathrm{a}} = F_i^{\mathrm{a}} + N_i(\boldsymbol{\xi}_k) F_k^{\mathrm{f}}. \tag{29}$$

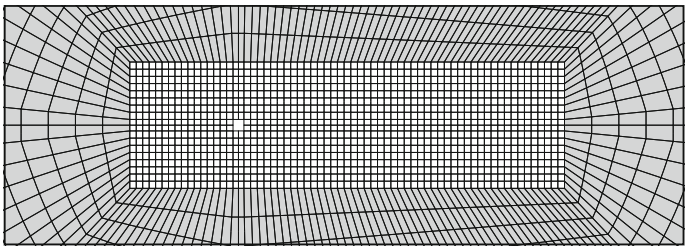


Fig. 15 Details of the conforming mesh. A 2D cut in the xy-plane is depicted

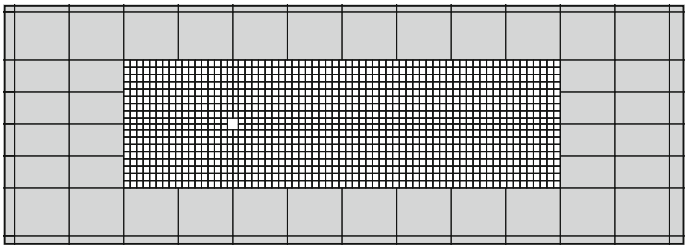


Fig. 16 Details of the non-conforming mesh. A 2D cut in the xy-plane is depicted

In order to demonstrate the capability of the non-conforming grid technique, a few test cases are defined. Cuts of the reference mesh and our non-conforming mesh are depicted in the vicinity of the source domain Ω_{a1} in Figs. 15 and 16. Thereby, the following different grids and order of FE basis functions have been investigated:

- **SQCONF**: A conforming mesh consisting of 20-node hexahedra is used, see Fig. 15. The results in Kaltenbacher et al. (2010) have been obtained with this mesh.
- **SQQUQU**: The same mesh as in the SQCONF case is used in Ω_{a1} (source domain) but a Cartesian 20-node hexahedra mesh is used in the propagation domain Ω_{a2} (cf. Fig. 16). The mesh in Ω_{a2} has a very fine discretization, namely, about 120 degrees of freedom per wavelength at 55 Hz.
- **SQLIQU**: For this case linear elements (8-node) of the mesh in Ω_{a1} are used, which contain the same corner nodal sources as the one in SQCONF. This reduces the number of unknowns in the source region by a factor of four compared to the quadratic mesh. The same 20-node hexahedra mesh as in SQQUQU is used in the propagation domain.
- **SQLILI**: In order to substantially decrease the number of unknowns also in the propagation region Ω_{a2} trilinear hexahedron elements are used in that domain. In comparison to the SQLIQU case, the topology of the mesh in Ω_{a2} stays the same. This cuts down the number of unknowns to one fourth also in the propagation domain. One can expect little or no impact on the accuracy of the solution, since the mesh still has a resolution of about 60 degrees of freedom per wavelength at

Table 3 Number of unknowns and wall clock times for the square cylinder cases

Test case	Ω_{a1}	Ω_{a2}	Total	Wall clock time (s)
SQCONF	93,781	444,652	538,433	1404.0
SQQUQU	93,781	142,163	236,437	202.0
SQLIQU	24,177	142,163	166,833	131.0
SQLILI	24,177	36,386	60,738	26.0

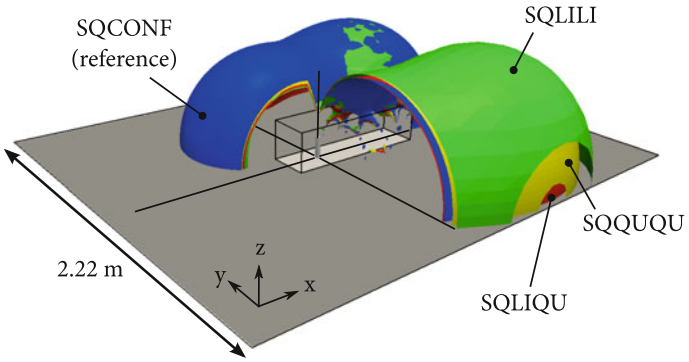


Fig. 17 Square cylinder with isosurfaces of acoustic pressure ($p_a = 6$ mPa)

55 Hz. Compared to the reference setup, the number of unknowns is reduced by a factor of nine.

Table 3 gives an overview of the number of unknowns for the acoustic computations. The solver and the number of threads on the used computer hardware are kept the same. The times clearly correspond to the total number of unknowns for each case. The results show that the computation in the SQLILI case is 54 times faster than in the reference case while still yielding comparable results. In all non-conforming cases, the Lagrange multiplier is defined on the coarse discretization of the surface Γ_I. In all simulations Ω_{a2} is used as the slave side and Ω_{a1} is used as the master.

The results of the computations for the vortex shedding frequency at 55 Hz are shown in Fig. 17 as isosurfaces of the acoustic pressure field and in Figs. 18, 19 and 20 as directivity plots in the xy-plane. It is obvious that all four configurations produce almost the same results.

5.2 Axial Fan

The cabin noise of modern ground vehicles is highly affected by flow related noise sources. This is especially the case, when the vehicle is not moving. Thereby, fan and outlet of air-conditioning systems are main acoustic sources and may reduce the

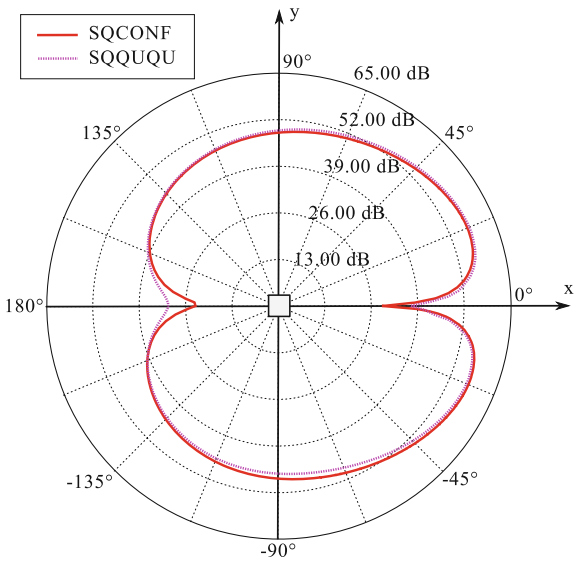


Fig. 18 Directivity plots of sound pressure levels at z = 0 in 1 m distance of the square cylinder for SQQUQU

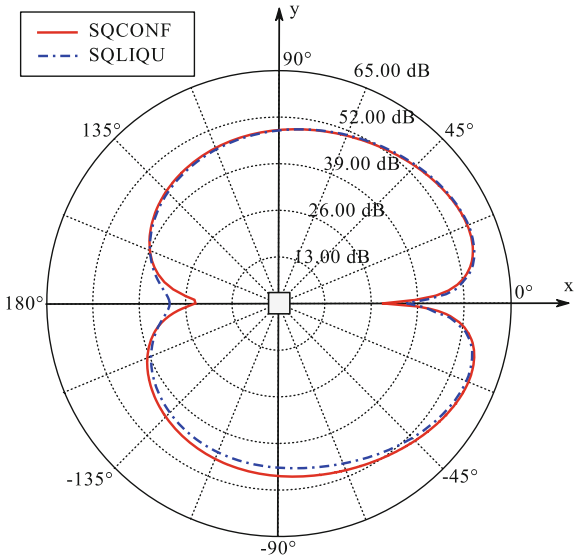


Fig. 19 Directivity plots of sound pressure levels at z = 0 in 1 m distance of the square cylinder for SQLIQU

comfort significantly. Rotating fans generate a highly turbulent flow field and can be identified as the main noise source in air conditioning units. Therefore, we focus on flow simulations of rotating fans in air conditioning units using the Arbitrary Mesh Interface (AMI) as implemented in OpenFOAM. For the computation of the acoustic sources, highly accurate unsteady CFD simulation data is needed. Therefore, the transient simulations are carried out by using a DES (Detached Eddy Simulation) turbulence model to accurately resolve the complex flow field. In addition, acoustic

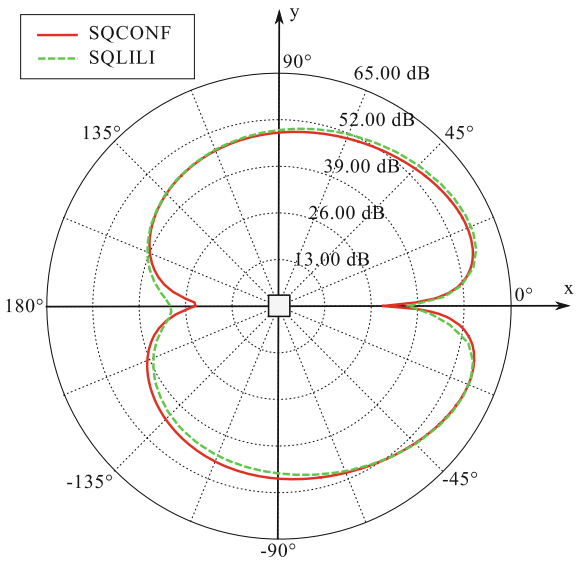


Fig. 20 Directivity plots of sound pressure levels at z = 0 in 1 m distance of the square cylinder for SQLILI

simulations applying the proposed Nitsche-type mortaring to couple the acoustic field between rotating and stationary sub-domains are performed.

Aeroacoustic Formulation The acoustic/viscous splitting technique for the prediction of flow induced sound was first introduced by Hardin and Pope (1994), and afterwards many groups presented alternative and improved formulations for linear and non linear wave propagation (Shen and Sørensen 1999; Ewert and Schröder 2003; Seo and Moon 2005; Munz et al. 2007). These formulations are all based on the idea, that the flow field quantities are split into compressible and incompressible parts. We apply a generic splitting of physical quantities to the Navier–Stokes equations. For this purpose we choose the following ansatz (Hüppe 2013)

$$p = \bar{p} + p_{\mathrm{ic}} + p_{\mathrm{c}} = \bar{p} + p_{\mathrm{ic}} + p_{\mathrm{a}} \tag{30}$$

$$\boldsymbol{v} = \bar{\boldsymbol{v}} + \boldsymbol{v}_{\mathrm{ic}} + \boldsymbol{v}_{\mathrm{c}} = \bar{\boldsymbol{v}} + \boldsymbol{v}_{\mathrm{ic}} + \boldsymbol{v}_{\mathrm{a}} \tag{31}$$

$$\rho = \bar{\rho} + \rho_{\mathrm{l}} + \rho_{\mathrm{a}}\,. \tag{32}$$

Thereby the field variables are split into mean ($\bar{p}$, $\bar{\boldsymbol{v}}$, $\bar{\rho}$) and fluctuating parts just like in the Linearized Euler Equations (LEE). In addition the fluctuating field variables are split into acoustic (p_{a}, $\boldsymbol{v}_{\mathrm{a}}$, ρ_{a}) and flow components (p_{ic}, $\boldsymbol{v}_{\mathrm{ic}}$). Finally, a density correction ρ_{l} is build in according to (32). This choice is motivated by the following assumptions:

- The acoustic field is a fluctuating field.
- The acoustic field is irrotational, i.e. $\nabla \times \boldsymbol{v}_{\mathrm{a}} = 0$, and therefore may be expressed by the acoustic scalar potential ψ_{a} via

$$\boldsymbol{v}_\mathrm{a} = -\nabla\psi_\mathrm{a}\,. \tag{33}$$

- The acoustic field requires compressible media and an incompressible pressure fluctuation is not equivalent to an acoustic pressure fluctuation.

By doing so, we arrive for an incompressible flow at the following perturbed convective wave equation (PCWE) (Kaltenbacher et al. 2016b)

$$\frac{1}{c^2}\,\frac{D^2\psi_\mathrm{a}}{Dt^2} - \Delta\psi_\mathrm{a} = -\frac{1}{c^2\bar{\rho}}\,\frac{Dp_\mathrm{ic}}{Dt}\,;\quad \frac{D}{Dt} = \frac{\partial}{\partial t} + \overline{\boldsymbol{v}}\cdot\nabla\,. \tag{34}$$

Now, as shown in Donea et al. (2004), we may apply an ALE (Arbitrary Lagrangian Eulerian) formulation to couple rotating and stationary domains. Thereby, our operator D/Dt changes to

$$\frac{D}{Dt} \rightarrow \frac{\tilde{D}}{\tilde{D}t} = \frac{\partial}{\partial t} + (\overline{\boldsymbol{v}} - \boldsymbol{v}_\mathrm{r})\cdot\nabla \tag{35}$$

with $\boldsymbol{v}_\mathrm{r}$ the mechanical velocity of rotating parts. Finally, the acoustic pressure p_a computes by

$$p_\mathrm{a} = \bar{\rho}\frac{\tilde{D}\psi_\mathrm{a}}{Dt}\,. \tag{36}$$

Thereby, PCWE is an exact reformulation of the acoustic perturbation equations (APE) (Ewert and Schröder 2003). This convective wave equation fully describes acoustic sources generated by incompressible flow structures and its wave propagation through flowing media. In addition, instead of the original unknowns acoustic pressure p_a and acoustic particle velocity $\boldsymbol{v}_\mathrm{a}$, this formulation has just the scalar unknown ψ_a, which strongly reduces computational time.

5.3 Numerical Computations

We investigate the aeroacoustic field of an axial fan in a duct as displayed in Fig. 21. The fan is embedded in a sound hard tube. The inlet and outlet openings on each side lead into a non reverberant chamber to emulate free field sound propagation. The rotational speed of the fan is about 1500 rpm, which results in a tip speed of the blades of 38.89 m/s. We use the OpenFOAM (Open Field Operation and Manipulation) CFD Toolbox version 2.3.0 for performing the flow computations. Since version 2.1.0 the arbitrary mesh interface (AMI) was implemented based on the algorithm described in Farrell and Maddison (2011). The AMI allows simulation across disconnected, but adjacent mesh domains, which are especially required for rotating geometries.

The flow solution is computed using an adapted version of the pimpleDyMFoam solver implemented in OpenFOAM, which can handle dynamic meshes with a time

Fig. 21 Axial fan

step size of $\Delta t = 10\,\mu$s. For the CFD computation a hex-dominant finite volume mesh consisting of 29.8 million cells was created by using the automatic mesh generator HEXPRESS^{TM}/Hybrid from Numeca. The transient simulation was carried out by using a detached-eddy simulation based on the Spalart–Allmaras turbulence model to accurately resolve the complex flow field (Spalart and Allmaras 1994). The calculation was performed on the Vienna Scientific Cluster VSC2 with 256 cores. Figure 22 displays the velocity field for a characteristic time step. Based on the computed instationary flow pressure, we display the surface contours of the acoustic sources (substantial derivative of the incompressible flow pressure, see (34)) in Fig. 23 for a characteristic time step. In accordance to the flow computation, the rotating domain is embedded into a quiescent propagation region (see Fig. 24). Furthermore, we add at the inflow and outflow boundaries of the CFD domain two additional regions, on which we apply an advanced *Perfectly-Matched-Layer* technique to effectively approximate acoustic free field conditions (Kaltenbacher et al. 2013). Figure 25 displays the computed power spectral density of the acoustic pressure and compares it to the measured one. Thereby, we display the smoothed measured spectra obtained from the 30 s recorded pressure signals as well as the individual spectra by just using measured data of 0.1 s (in gray). The spectra based on our numerical simulation is computed out of a real time simulation of 0.06 s.

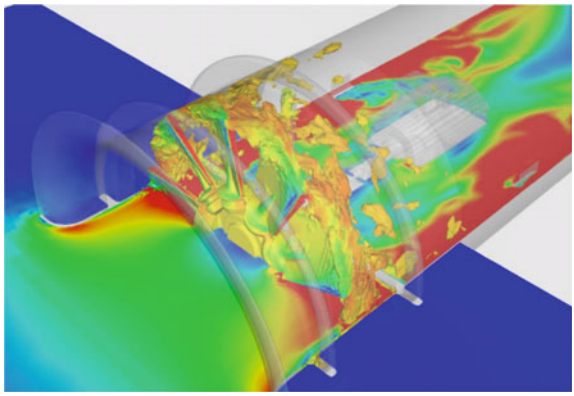


Fig. 22 Flow structure of a characteristic time step

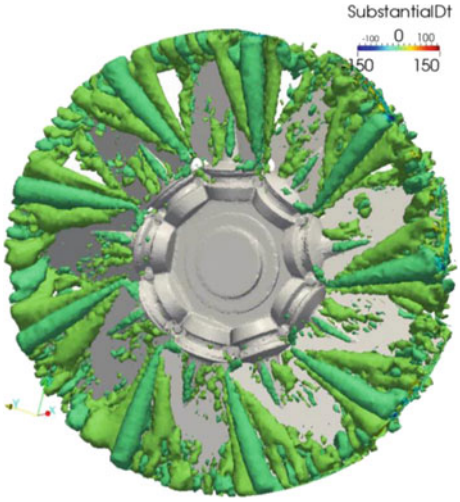


Fig. 23 Visualization of the acoustic source terms at a characteristic time step

5.4 Human Phonation

The voice production mechanisms have been investigated both by means of measurements (on physical replicas, excised animal or human larynges or in living subjects) and numerical simulations. The experimental investigation, especially in vivo, brings numerous complications. Since the advent of affordable high-performance computing, the computer simulation methods based on modeling the fundamental physical phenomena using partial differential equations and solving them numerically have been steadily gaining importance.

An extensive review of numerical models of human phonation can be found in Alipour et al. (2011). The vibration of the real human vocal folds is flow-induced. However, the fully coupled fluid-structure simulations, e.g., Link et al. (2009), Seo and Mittal (2011) always suffer from a lack of accurate geometrical and material

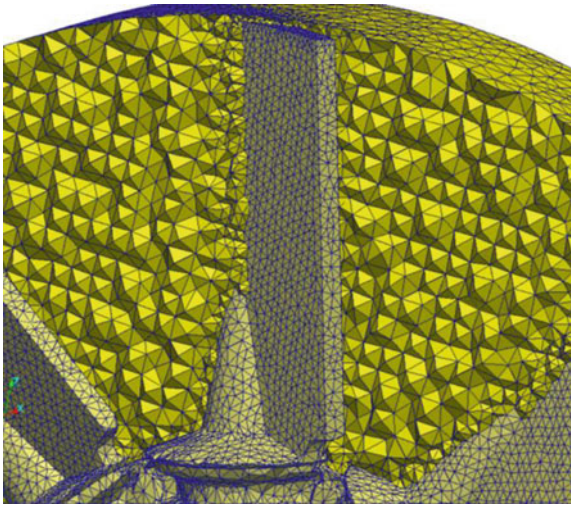


Fig. 24 Detail of the computational CFD grid

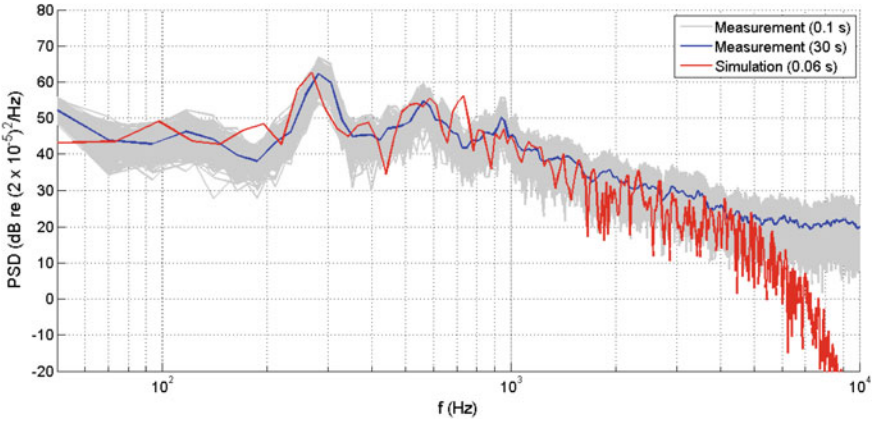


Fig. 25 Power spectral density of the acoustic pressure at measurement position

properties of the living tissues. This is due to the fact that the parameters are highly subject-specific, and also because most of the vocal fold tissue measurements, e.g., Zörner et al. (2010), Kelleher et al. (2013), are still hardly applicable in vivo to precisely determine the vocal fold material parameters. As shown by Zörner et al. (2013), the full fluid-structure interaction solution can be approximated by prescribed vocal fold motion, provided that the boundary conditions are set properly. In this case, it is crucial that the vocal fold vibration patterns mimic the motion of the real human vocal folds sufficiently well. The kinematic parameters have been intensively studied by videokymographic (Svec and Schutte 2012) and high-speed imaging methods (Döllinger et al. 2011), and the results of the experimental studies will be used in our model.

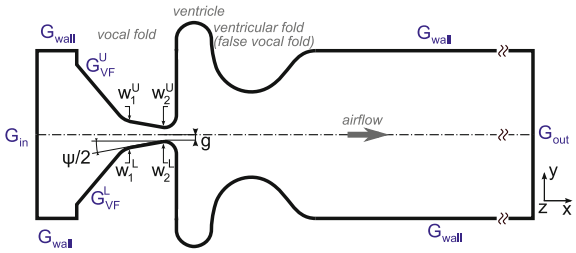


Fig. 26 Geometric model of the human larynx in coronal section. The length of the supra-glottal channel is not to scale

Table 4 Parameters of the kinematic model: Superscripts L and U refer to the lower and upper vocal fold, respectively, $D = 12\,\text{mm}$ is the anterior-posterior length of the vocal folds

$A_1{}^U$, $A_2{}^U$	$A_1{}^L$, $A_2{}^L$	ξ	$m(z)$
0.3 mm	0.3 mm	$\pi/2$	$\sin(\pi z/D)$

$\psi^U/2$	$\psi^L/2$	$g_{\max}$	$g_{\min}$
$-12° \ldots 12°$	$-12° \ldots 12°$	0.72 mm	0.2 mm

Geometry and vocal fold kinematics The flow field is solved on a simplified model of larynx, consisting of a short straight sub-glottal region, the vocal folds, ventricles and false vocal folds (FVFs), and a supra-glottal region (see Fig. 26). In the straight sub-glottal and supra-glottal segments, the model has a square cross-section of $12 \times 12\,\text{mm}$, with vocal and ventricular folds having a length of 7.2 and 6.3 mm, respectively. The detailed dimensions of the flow domain can be found in Šidlof et al. (2014). During the CFD simulation the vocal folds, forming part of the channel boundary, oscillate. The kinematics of the vocal folds were programmed to allow for two-degree-of-freedom, convergent-divergent motion of each of the vocal folds, with prescribed sinusoidal displacement of the inferior and superior vocal fold margins in the medial-lateral direction

$$\begin{aligned} w_1(z,t) &= w_{10} + A_1\,(1 - m(z)) + m(z) A_1 \sin(2\pi f t + \xi) \\ w_2(z,t) &= w_{20} + A_2\,(1 - m(z)) + m(z) A_2 \sin(2\pi f t)\,. \end{aligned} \tag{37}$$

In (37) f is the frequency of vibration, ξ the phase difference between the inferior and superior margin, $A_{1/2}$ the amplitudes and $m(z)$ the anterior-posterior modulation function determining the glottal opening shape (see also Fig. 26 and Table 4). The coordinates in (37) determine uniquely the glottal half-gap g and the medial surface convergence angle ψ.

Flow model and boundary conditions In regular human phonation the air flows at low Mach numbers (Ma < 0.2) and can thereby be regarded as incompressible. This sets the fluid density ρ to a constant value and results in the 3D time dependent incompressible Navier–Stokes equations. The frequency of vocal fold vibration is set to $f = 100\,\text{Hz}$, corresponding to Strouhal number in the order of $\text{St} = 0.001$. The airflow is driven by a pressure gradient, which mimics physiological conditions

Table 5 Boundary conditions for the velocity $\boldsymbol{v}$ and kinematic pressure $P = p/\rho$. Vector $\boldsymbol{n}$ denotes the unit outer normal, $\boldsymbol{u}$ is the prescribed boundary displacement

$\boldsymbol{v}$ [m/s]			P [m^2/s^2]
G_{in}	evaluated	$(\boldsymbol{v} \cdot \boldsymbol{n} < 0)$	$P + \frac{1}{2}\|\boldsymbol{v}\|^2 = 300$
	$\boldsymbol{v} = 0$	$(\boldsymbol{v} \cdot \boldsymbol{n} > 0)$	
G_{out}	$\partial \boldsymbol{v}/\partial \boldsymbol{n} = 0$	$(\boldsymbol{v} \cdot \boldsymbol{n} > 0)$	$P = 0$
	$\boldsymbol{v} = 0$	$(\boldsymbol{v} \cdot \boldsymbol{n} < 0)$	
G_{VF}^{L}	$\boldsymbol{v} = \frac{\partial \boldsymbol{u}}{\partial t}$		$\frac{\partial P}{\partial \boldsymbol{n}} = 0$
G_{VF}^{U}	$\boldsymbol{v} = \frac{\partial \boldsymbol{u}}{\partial t}$		$\frac{\partial P}{\partial \boldsymbol{n}} = 0$
G_{wall}	$\boldsymbol{v} = 0$		$\frac{\partial P}{\partial \boldsymbol{n}} = 0$

with a constant lung pressure at the inlet and a zero relative pressure at the outlet. The boundary conditions for the velocity $\boldsymbol{v}$ and kinematic pressure $P = p/\rho$ are summarized in Table 5. The Navier–Stokes equations were discretized using a collocated cell-centered variant of the finite volume method for unstructured meshes. The numerical solution was implemented with the help of OpenFOAM. The discretization scheme for the time derivative is first-order Euler implicit, a total variation diminishing (TVD) scheme for the convection term and central differencing with explicit non-orthogonal correction for the diffusion term. The time step Δt is adjusted automatically during the transient solution so that the maximum local Courant number is kept below a predefined limit. In the current simulations, the Courant number was kept below 1, resulting in a time step size Δt of $5 \cdot 10^{-7}$ s–$1.5 \cdot 10^{-6}$ s. The discretized Navier–Stokes equations were solved by a segregated solver based on a modified pressure implicit with splitting of operators (PISO) algorithm (Issa 1986), with the preconditioned biconjugate gradient linear solver for the momentum equations and algebraic multigrid for the pressure equation.

CFD results The results of the CFD simulations are displayed in Fig. 27 in mid-coronal z-normal sections at four time instants corresponding to the closing phase, maximum closure, opening phase and maximum opening. The velocity fields are taken from the 19th period of vibration, when the flow is already fully developed. Figure 28 shows the velocity magnitude in the transverse planes and the jet contours (velocity isosurfaces), giving insight in the three-dimensionality of the supra-glottal flow fields. The CFD results confirm the experimental findings in Triep and Brücker (2010) and numerical simulations in Schwarze et al. (2011), who showed that this geometry promotes the phenomenon of jet axis switching. The jet, mostly planar and aligned along the anterior-posterior axis within glottis, changes its orientation further downstream of the glottis and aligns in the medial-lateral direction in the second half of the opening phase and first half of the closing phase. The axis switching also induces complex 3D vortex structures within the ventricles.

Acoustic model The acoustic domain consists of three parts: The first part is the larynx, which contains the aeroacoustic sources and corresponds to the flow domain. Attached to it is the second part, the vocal tract, which is a 18.25 cm long tube with

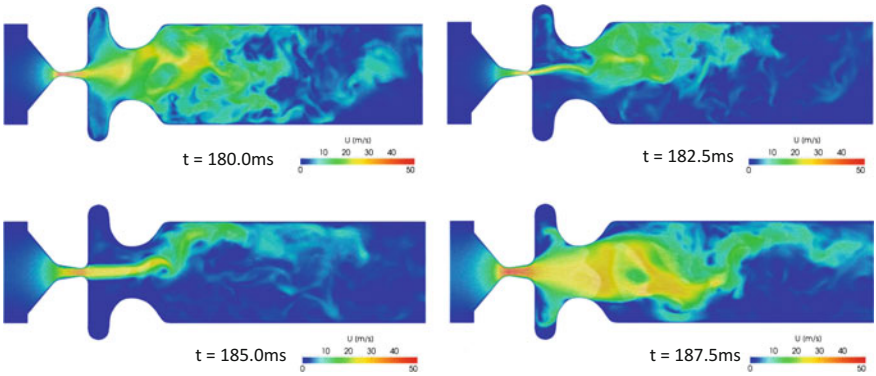


Fig. 27 Velocity magnitude in z-normal (coronal) mid-plane at four time instants during the 19th period of oscillation. From *left* to *right* closing phase, maximum closure, opening phase, maximum opening

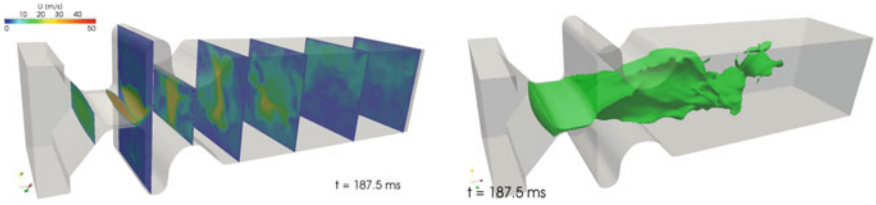


Fig. 28 Velocity magnitude and jet contours (isosurface at $u = 15$ m/s), maximum opening

varying diameter along the center axis. The vocal tract acts as an acoustic filter and modulates the generated sound, by amplifying or reducing the amplitudes at certain frequencies. For this purpose the vocal tract geometry representing the vowel /u/ ("who") was chosen. Exact dimensions were taken from Story et al. (1996), where 18 three-dimensional vocal tract shapes were acquired by means of magnetic resonance imaging (MRI).

The last part of the acoustic domain is the propagation region, a $2.5 \times 2.5 \times 2.5\,\text{cm}^3$ big box, which is added at the end of the vocal tract. Its purpose is to capture the sound wave in 1 cm distance from the mouth at the monitoring position "MIC". In Fig. 29 the geometric model used for the acoustic simulation together with the monitoring point is plotted.

The grid size of the acoustic simulation can be chosen considerably coarser than the characteristic length of the CFD grid (0.15 mm). Therefore the acoustic mesh is composed of hexahedron elements with a characteristic length of 0.2 mm inside the glottis region (corresponds to the CFD domain). The non-conforming grid technique allows us to directly connect the source (flow domain) and propagation domain (corresponds to the vocal tract). The overall grid for the acoustic simulation is fine enough for computations up to 3.5 kHz. All channel walls are considered to be fully reflecting, and perfectly matched layers (PML) are located at the inflow (1 cm in front

of the glottis) and surrounding the propagation region (see Fig. 29). The PML at the inlet is 5 mm long in x-direction and 6 mm in normal direction to the propagation region.

Acoustics results The acoustic field is computed by the FE method solving the perturbed convective wave equation (PCWE) as described in Sect. 5.2. For the acoustic source analysis, the substantial derivative of the incompressible pressure p_{ic} is Fourier transformed over the whole domain. The fundamental frequency is found at 100 Hz, as this is the frequency the vocal folds are being driven. Investigating the acoustic sources show that the main contributions are inside the glottis, as the contour plots in Fig. 30 reveal. Studying the source distribution for a non-harmonic frequency of 2665 Hz, which is a random representative of the broadband spectrum, reveals that the sources are distributed downstream which correlates to the vortex shedding region (see Fig. 28). For other frequencies of the broadband spectrum, the results are comparable, concerning distribution and amplitude of the source region.

The monitoring point "MIC" is situated 1 cm downstream of the mouth, as sketched in Fig. 29. The computed acoustic spectrum is plotted in Fig. 31 and shows

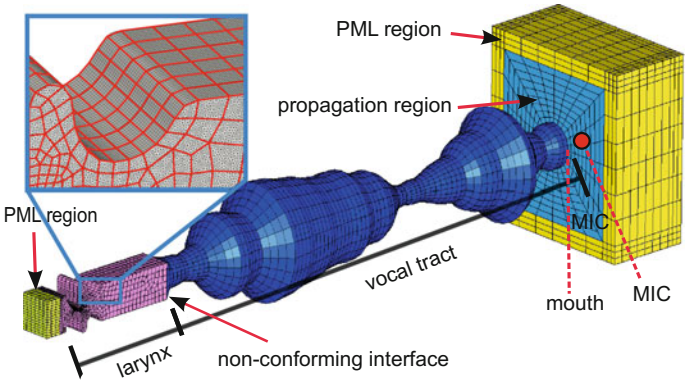


Fig. 29 Geometry and mesh for the acoustic simulation. Larynx, vocal tract, propagation region, perfectly matched layer (PML) regions and comparison of the fine CFD grid and coarse acoustic grid

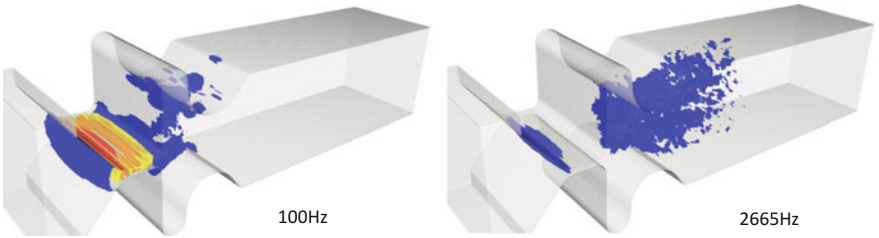


Fig. 30 Acoustic sources at main frequency (100 Hz) and at 2665 Hz

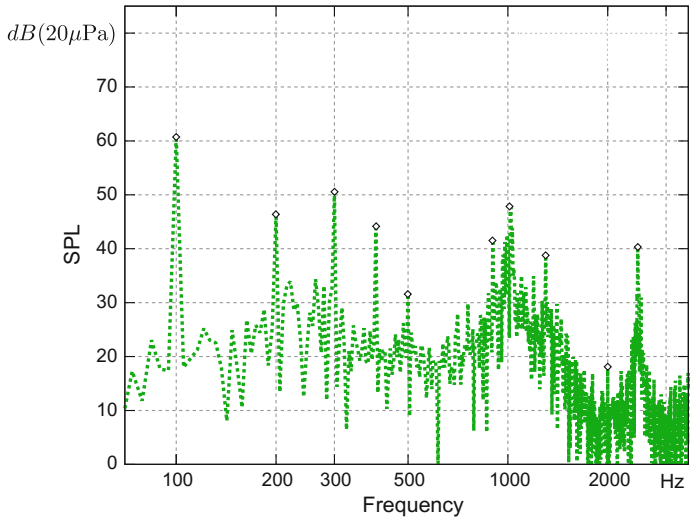


Fig. 31 Acoustic sound spectra at a monitoring point "MIC" for the vocal tract model /u/. Harmonics are emphasized with the symbols ◇

that the first harmonic has the largest amplitude, and all other harmonics are up to 15 dB lower. Furthermore, the amplitudes at non-harmonics are consistently smaller by about 5 dB.

Acknowledgements The author wishes to acknowledge his former Ph.D. students Andreas Hüppe, Simon Triebenbacher and Stefan Zörner for main contributions towards non-conforming grid techniques and its applications. Furthermore, I wish to tanks my colleague Barbara Wohlmuth (Technische Universität München, Germany) for our longtime cooperation on non-conforming grid techniques. Finally, many thanks to Stefan Schoder for proof reading and his usefull suggestions.

References

Alipour, F., Brücker, C., Cook, D., Gömmel, A., Kaltenbacher, M., Willy, M., et al. (2011). Mathematical models and numerical schemes for the simulation of human phonation. *Current Bioinformatics*, *6*(3), 323–343.

Becker, S., Hahn, C., Kaltenbacher, M., & Lerch, R. (2008). Flow-induced sound of wall-mounted cylinders with different geometries. *AIAA Journal*, *46*(9), 2265–2281.

Bernardi, C., Maday, Y., & Patera, A. T. (1994). A new nonconforming approach to domain decomposition: The mortar element method. *Nonlinear partial differential equations and their applications. Collège de France Seminar, Vol. XI (Paris, 1989–1991)* (Vol. 299, pp. 13–51). Pitman research notes in mathematics series. Harlow: Longman Scientific and Technical.

Dokeva, N. (2006). Scalable mortar methods for elliptic problems on many subregions. Ph.D. thesis, University of Southern California.

Döllinger, M., Kobler, J., Berry, A., Mehta, D., Luegmair, G., & Bohr, C. (2011). Experiments on analysing voice production: Excised (human, animal) and in vivo (animal) approaches. *Current Bioinformatics*, *6*(3), 286–304.

Donea, J., Huerta, A., Ponthot, J. P., & Rodriguez-Ferran, A. (2004). Arbitrary Lagrangian-Eulerian methods. *Encyclopedia of computational mechanics*. New York: Wiley.

Durst, F., & Schäfer, M. (1996). A parallel block-structured multigrid method for the prediction of incompressible flows. *International Journal for Numerical Methods in Fluids*, *22*, 549–565.

Escobar, M. (2007). Simulation of flow induced noise. Ph.D. thesis, Department of Sensor Technology, University of Erlangen-Nuremberg.

Ewert, R., & Schröder, W. (2003). Acoustic perturbation equations based on flow decomposition via source filtering. *Journal of Computational Physics*, *188*, 365–398.

Farrell, P. E., & Maddison, J. R. (2011). Conservative interpolation between volume meshes by local Galerkin projection. *Computer Methods in Applied Mechanics and Engineering*, *200*, 89–100.

Flemisch, B., Kaltenbacher, M., Triebenbacher, S., & Wohlmuth, B. (2012). Non-matching grids for a flexible discretization in computational acoustics. *Communications in Computational Physics*, *11*(2), 472–488.

Fritz, A., Hüeber, S., & Wohlmuth, B. (2004). A comparison of mortar and Nitsche techniques for linear elasticity. In *CALCOLO*.

Gander, M. J., & Japhet, C. (2009). An algorithm for non-matching grid projections with linear complexity. *Domain decomposition methods in science and engineering XVIII*. Berlin: Springer.

Grabinger, J. (2007). Mechanical-acoustic coupling on non-matching finite element grids. Master's thesis, University Erlangen-Nuremberg.

Greiner, G., & Hormann, K. (1998). Efficient clipping of arbitrary polygons. *ACM Transactions on Graphics*, *17*, 71–83.

Hahn, C. (2008). Experimentelle Analyse und Reduktion aeroakustischer Schallquellen an einfachen Modellstrukturen. Ph.D. thesis, University of Erlangen-Nuremberg.

Hansbo, A., Hansbo, P., & Larson, M. G. (2003). A finite element method on composite grids based on Nitsche's method. *ESAIM: Mathematical Modelling and Numerical Analysis*, *37*(3), 495–514.

Hardin, J. C., & Pope, D. S. (1994). An acoustic/viscous splitting technique for computational aeroacoustics. *Theoretical and Computational Fluid Dynamics*, *6*, 323–340.

Heinstein, M. W., & Laursen, T. A. (2003). Consistent mesh tying methods for topologically distinct discretized surfaces in non-linear solid mechanics. *International Journal for Numerical Methods in Engineering*, *57*, 1197–1242.

Hughes, T. J. R. (2000). *The finite element method*. New York: Dover.

Hüppe, A. (2013). Spectral finite elements for acoustic field computation. Ph.D. thesis, University of Klagenfurt, Austria.

Issa, R. I. (1986). Solution of the implicitly discretised fluid flow equations by operator-splitting. *Journal of Computational Physics*, *62*(1), 40–65.

Kaltenbacher, B., Kaltenbacher, M., & Sim, I. (2013). A modified and stable version of a perfectly matched layer technique for the 3-d second order wave equation in time domain with an application to aeroacoustics. *Journal of Computational Physics*, *235*, 407–422.

Kaltenbacher, M. (2015). *Numerical simulation of mechatronic sensors and actuators - finite elements for computational multiphysics* (3rd ed.). Berlin: Springer.

Kaltenbacher, M., Escobar, M., Ali, I., & Becker, S. (2010). Numerical simulation of flow-induced noise using LES/SAS and Lighthill's acoustics analogy. *International Journal for Numerical Methods in Fluids*, *63*(9), 1103–1122.

Kaltenbacher, M., Hüppe, A., Grabinger, J., & Wohlmuth, B. (2016a). Modeling and finite element formulation for acoustic problems including rotating domains. *AIAA Journal*, *54*, 3768–3777.

Kaltenbacher, M., Hüppe, A., Reppenhagen, A., Tautz, M., Becker, S., & Kühnel, W. (2016b). Computational aeroacoustics for HVAC systems utilizing a hybrid approach. *SAE International*, *9*.

Kelleher, J. E., Siegmund, T., Du, M., Naseri, E., & Chan, R. W. (2013). Empirical measurements of biomechanical anisotropy of the human vocal fold lamina propria. *Biomechanics and Modeling in Mechanobiology*, *12*(3), 555–567.

Klöppel, T., Popp, A., Küttler, U., & Wall, W. (2011). Fluid - structure interaction for non-conforming interfaces based on a dual morta formulation. *Computer Methods in Applied Mechanics and Engineering*, *200*(45), 3111–3126.

Köck, H., Eisner, S., & Kaltenbacher, M. (2015). Electrothermal multiscale modeling and simulation concepts for power electronics. *IEEE Transactions on Power Electronics*, *99*.

Langer, U., & Steinbach, P. (2003). Boundary element tearing and interconnecting methods. *Computing*, *71*(3), 205–228.

Link, G., Kaltenbacher, M., Breuer, M., & Dllinger, M. (2009). A 2d finite-element scheme for fluidsolidacoustic interactions and its application to human phonation. *Computer Methods in Applied Mechanics and Engineering*, *198*, 3321–3334.

Menter, F., & Egorov, Y. (2005). A scale adaptive simulation model using two-equation models. In *43rd AIAA Aerospace Sciences Meeting and Exhibit*, number AIAA-2005-1095.

Munz, C. D., Dumbser, M., & Roller, S. (2007). Linearized acoustic perturbation equations for low Mach number flow with variable density and temperature. *Journal of Computational Physics*, *224*, 352–364.

Park, K. C., & Felippa, C. A. (2002). A simple algorithm for localized construction of non-matching structural interfaces. *International Journal for Numerical Methods in Engineering*, *53*, 2117–2142.

Puso, M. A. (2004). A 3d mortar method for solid mechanics. *International Journal for Numerical Methods in Engineering*, *59*, 315–336.

Puso, M. A., & Laursen, T. A. (2002). A 3d contact smoothing method using Gregory patches. *International Journal for Numerical Methods in Engineering*, *54*, 1161–1194.

Schwarze, R., Mattheus, W., Klostermann, J., & Brücker, C. (2011). Starting jet flows in a three-dimensional channel with larynx-shaped constriction. *Computers and Fluids*, *48*(1), 68–83.

Seo, J. H., & Mittal, R. (2011). A high-order immersed boundary method for acoustic wave scattering and low-Mach number flow-induced sound in complex geometries. *Journal of Computational Physics*, *230*(4), 1000–1019.

Seo, J. H., & Moon, Y. J. (2005). Perturbed compressible equations for aeroacoustic noise prediction at low Mach numbers. *AIAA Journal*, *43*, 1716–1724.

Shen, W. Z., & Sørensen, J. N. (1999). Aeroacoustic modelling of low-speed flows. *Theoretical and Computational Fluid Dynamics*, *13*, 271–289.

Šidlof, P., Zörner, S., & Hüppe, A. (2014). A hybrid approach to the computational aeroacoustics of human voice production. *Biomechanics and Modeling in Mechanobiology*, 1–16.

Spalart, P. R., & Allmaras, S. R. (1994). A one-equation turbulence model for aerodynamic flows. *Recherche Aerospatiale*, *1*, 5–21.

Story, B. H., Titze, I. R., & Hoffman, E. A. (1996). Vocal tract area functions from magnetic resonance imaging. *The Journal of the Acoustical Society of America*, *100*(1), 537–554.

Svec, J., & Schutte, H. K. (2012). Kymographic imaging of laryngeal vibrations. *Current Opinion in Otolaryngology and Head and Neck Surgery*, *20*(6), 458–465.

Triep, M., & Brücker, C. (2010). Three-dimensional nature of the glottal jet. *The Journal of the Acoustical Society of America*, *127*(3), 1537–1547.

Wohlmuth, B. I. (2000). A mortar finite element method using dual spaces for the Lagrange multiplier. *SIAM Journal on Numerical Analysis*, *38*(3), 989–1012.

Zörner, S., Kaltenbacher, M., Lerch, R., Sutor, A., & Döllinger, M. (2010). Measurement of the elasticity modulus of soft tissues. *Journal of Biomechanics*, *43*(8), 1540–1545.

Zörner, S., Kaltenbacher, M., & Döllinger, M. (2013). Investigation of prescribed movement in fluid-structure interaction simulation for the human phonation process. *Computers and Fluids*, *86*, 133–140.

Boundary Element Method for Time-Harmonic Acoustic Problems

Steffen Marburg

Abstract This chapter presents an introduction to the solution of time-harmonic acoustic problems by a boundary element method (BEM). Specifically, the Helmholtz equation with admittance boundary conditions is solved in 3d. The chapter starts with a derivation of the Kirchhoff–Helmholtz integral equation from a residual formulation of the Helmholtz equation. The discretization process with introduction of basis and test functions is described and shown for the collocation and the Galerkin method. Thereafter, only collocation is used. The next section describes the application of field sources and incident wave fields on behalf of a particular solution. This is followed by a discussion on field point evaluation and a detailed description on the evaluation of the system matrix entries. The latter starts with the integral free terms, continues with an adaptive integration strategy for regular and quasi-singular integrals and finishes with an integration strategy for singular integrals. Subsequent sections discuss the choice of boundary elements and the methods to deal with the well-known non-uniqueness problem in BEM. While it has become obvious for the former problem that discontinuous Lagrangian elements perform the best, in the latter case the author is convinced that the Burton and Miller method is the only safe and efficient choice to avoid irregular frequencies. The next subsection explains the motivation for and the basic idea of fast boundary element techniques and it concludes with a discussion about the cases when these fast techniques are actually reasonable. A section on structure fluid interaction is not just describing the so-called mortar formulation but also shows that a (non-local) boundary admittance may contain the complete information about the interaction between fluid and structure. The final two subsections deal with symmetric and periodic problems on the one side and with panel contribution analysis on the other side. Throughout this chapter, numerous different examples are presented. In some cases, the author chose simple one-dimensional examples which may be solved analytically. Other examples are rather industrial applications such as sedan cabin compartments, diesel engine radiation, a tire noise problem and the computation of common room acoustic measures for a music recording studio.

S. Marburg (✉)
Vibroacoustics of Vehicles and Machines, Technical University of Munich,
Boltzmannstraße 15, 85748 Garching Bei München, Germany
e-mail: steffen.marburg@tum.de

M. Kaltenbacher (ed.), *Computational Acoustics*, CISM International Centre
for Mechanical Sciences 579, DOI 10.1007/978-3-319-59038-7_3

1 Introduction

There is a wide range of papers about boundary element methods in acoustics. They are quite difficult to survey. While the boundary integral equations have already been developed in the 19th century, numerical methods for solution of these equations began to be developed in the early second half of the 20th century. A few textbooks on boundary element methods in acoustics were published over the last 30 years starting with the edition by Ciskowski and Brebbia (1991), followed by a monograph by Kirkup (1998) and two editions by Wu (2000b) and von Estorff (2000). In this context, the edition on finite and boundary element methods by Marburg and Nolte (2008a) and the interesting dissertation (published as a book) by do Rego Silva (1993) are mentioned as well.

It is the purpose of this chapter to present some basic formulations of boundary element methods (BEM) for linear time-harmonic acoustic problems. This includes many computational aspects such as numerical integration and choice of boundary elements. Furthermore, it includes discussion of the well-known non-uniqueness problem in BEM, techniques for structure-fluid interaction and others.

Within this chapter, problems are defined in a domain Ω which is bounded by the closed boundary Γ. Ω_c is the complementary domain and $\vec{\boldsymbol{n}}$ denotes the normal vector pointing into Ω_c. Interior and exterior problems are distinguished. Interior domains consider a finite domain Ω, cf. left subfigure in Fig. 1 while exterior problems consider domains Ω of infinite extent but finite complementary domain Ω_c, cf. right subfigure in Fig. 1.

It has been shown in the previous chapter that acoustic problems are governed by the linear wave equation as

$$\Delta \tilde{p}(\vec{\boldsymbol{x}}, t) = \frac{1}{c^2} \frac{\partial^2 \tilde{p}(\vec{\boldsymbol{x}}, t)}{\partial t^2} \qquad \vec{\boldsymbol{x}} \in \Omega \subset \mathbb{R}^d. \tag{1}$$

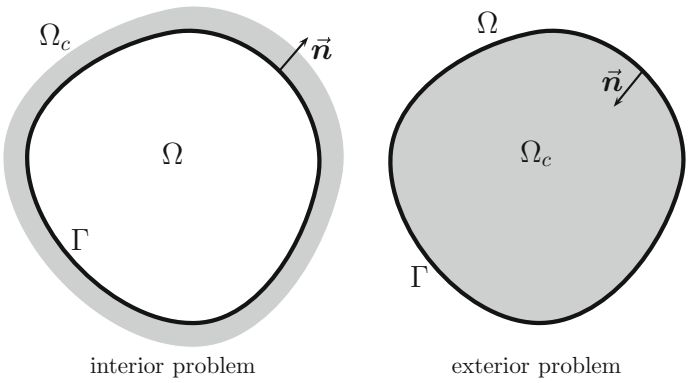


Fig. 1 Definition of regions Ω and Ω_c, boundary Γ and outward normal vector $\vec{\boldsymbol{n}}$

This equation is valid for the sound pressure $\tilde{p}$ depending on position $\vec{x}$ and time t whereas c is the speed of sound. The space dimension d is three in real applications, but can be two or one in certain cases. To obtain a solution, the differential equation requires boundary conditions and initial conditions, which will be specified when used.

For time-harmonic problems, a time dependence is introduced. Herein, we use the harmonic time-dependence

$$\tilde{p}(\vec{x}, t) = \Re\left\{p(\vec{x})\, \mathrm{e}^{-\mathrm{i}\omega t}\right\}. \tag{2}$$

Harmonic time dependence with the angular frequency ω could also be defined with a positive exponent of the exponential function. Both are possible but further steps must be adjusted to this choice. The negative exponent is chosen since it ensures that radiated waves are actually traveling outward whereas a positive exponent indicates an inward traveling wave. However, this difference is more formal than of scientific relevance. In literature, Eq. (2) is often given without the limitation to the real part on the right hand side. Although well-known what is meant here, this would be a wrong representation since the time-dependent physical sound pressure $\tilde{p}(\vec{x}, t)$ is a real valued quantity. The product of the complex valued time-harmonic sound pressure $p(\vec{x})$ and the complex valued exponential function, however, will be complex in general. According to redundant information on the right hand-side of Eq. (2), the user can either choose the real part or the imaginary part of this product.

Applying the time-harmonic dependence of p to Eq. (1) leads to the Helmholtz-equation, or harmonic wave equation, for the sound pressure

$$\Delta p(\vec{x}) + k^2 p(\vec{x}) = 0 \qquad \vec{x} \in \Omega \subset \mathbb{R}^d\,. \tag{3}$$

where the wavenumber $k = \omega/c$ is introduced. In most cases, admittance boundary conditions are assumed. They are equivalent to Robin boundary conditions which may degenerate to Neumann boundary conditions if the admittance is zero. This condition is written as

$$v_f(\vec{x}) - v_s(\vec{x}) = Y(\vec{x})\, p(\vec{x}) \qquad \vec{x} \in \Gamma \subset \mathbb{R}^{d-1}\,. \tag{4}$$

Y represents the boundary admittance which relates the sound pressure to the difference between the normal components of the fluid particle velocity v_f and the underlying structural particle velocity v_s. The normal component of the fluid particle velocity is related to the normal derivative of the sound pressure p by means of the linearized Euler equation in frequency domain as

$$v_f(\vec{x}) = \frac{1}{a}\frac{\partial p(\vec{x})}{\partial n(\vec{x})} = \frac{1}{\mathrm{i}\omega\varrho_0}\frac{\partial p(\vec{x})}{\partial n(\vec{x})} \tag{5}$$

where ϱ_0 represents the ambient density of the fluid. Note that for time dependence $\mathrm{e}^{\mathrm{i}\omega t}$, the constant a takes the conjugate value $a = -\mathrm{i}\omega\varrho_0$.

In some cases, it is useful to consider the Dirichlet boundary conditions. The Robin condition as formulated in Eq. (4) is not suited for this case. Instead, we may use the Robin condition as an impedance boundary condition with the impedance $Z(\vec{x})$ as

$$Z(\vec{x})\left[v_f(\vec{x}) - v_s(\vec{x})\right] = p(\vec{x}) \qquad \text{and} \qquad Z(\vec{x}) = \frac{1}{Y(\vec{x})} \,. \tag{6}$$

In case of a homogeneous Dirichlet boundary condition, the value of the impedance is zero and, thus, leads to $p(\vec{x}) = 0$. Obviously, the inhomogeneous Dirichlet condition results in $p(\vec{x}) = p_0(\vec{x})$.

The boundary value problem assumes a locally acting boundary admittance relating particle velocities of the fluid and the underlying structure to the sound pressure. For the several different problems considered in this chapter, the author considers simplified boundary conditions where either $v_s = 0$ or $Y = 0$.

In addition to fulfilling the Helmholtz equation and its boundary conditions, solutions of external problems require fulfillment of the decay condition at infinity, i.e. the Sommerfeld radiation condition. This is formulated in two steps for the sound pressure as

$$\begin{aligned} p &= O(r^{-\alpha}) \qquad \text{and} \\ \frac{\partial p}{\partial r} &- \mathrm{i}kp = o(r^{-\alpha}) \qquad \text{for} \quad r \to \infty \,, \end{aligned} \tag{7}$$

with $\alpha = (d-1)/2$ and r denoting the distance between an arbitrary field point and a point close to a source. Hence, the first expression of Eq. (7) formulates the decay rate of the solution of the Helmholtz equation, i.e. the sound pressure p, whereas the second expression requires the left hand-side to decay faster than $r^{-(d-1)/2}$. A valuable description in a rigorous form is given in the book by Ihlenburg (1998). Clearly, the Sommerfeld condition is a decay condition only for $d > 1$. Note that the minus sign on the left hand-side of the second part of the Sommerfeld condition changes into a plus sign if the time dependence is chosen to be $\mathrm{e}^{\mathrm{i}\omega t}$.

There are many practical problems with speed of sound and fluid density depending on position as $c = c(\vec{x})$ and $\varrho_0 = \varrho_0(\vec{x})$, e.g. in underwater acoustics and in atmospheric sound propagation. However, for the sake of simplicity, these cases will not be considered here.

2 Derivation of the Boundary Element Formulation

2.1 Weak Formulation

The boundary element formulation is based on a weighted residual formulation of the Helmholtz equation (3). It is the same for the finite element formulation. The following derivation is rather similar to the one by Marburg and Nolte (2008b) which shows that both, FEM and BEM, stem from the same origin.

A weighted residual formulation is based on introducing the weight function $\chi(\vec{x})$ and "testing" it with the Helmholtz operator such that

$$\int_\Omega \chi(\vec{x}) \left[\Delta p(\vec{x}) + k^2 p(\vec{x})\right] d\Omega(\vec{x}) \;=\; 0\,. \tag{8}$$

Integrating by parts gives

$$\begin{aligned}\int_\Omega \chi(\vec{x}) \left[\Delta p(\vec{x}) + k^2 p(\vec{x})\right] d\Omega(\vec{x}) \;=\; \\ \int_\Gamma \chi(\vec{x}) a v_f(\vec{x}) d\Gamma(\vec{x}) \;-\; \int_\Omega \left[\vec{\nabla}\chi(\vec{x}) \cdot \vec{\nabla} p(\vec{x}) - k^2 \chi(\vec{x}) p(\vec{x})\right] d\Omega(\vec{x}) \;=\; 0\end{aligned} \tag{9}$$

and then

$$\begin{aligned}\int_\Omega \chi(\vec{x}) \left[\Delta p(\vec{x}) + k^2 p(\vec{x})\right] d\Omega(\vec{x}) \;=\; \int_\Gamma \chi(\vec{x}) a v_f(\vec{x}) d\Gamma(\vec{x}) \;+ \\ -\; \int_\Gamma \frac{\partial \chi(\vec{x})}{\partial n(\vec{x})} p(\vec{x}) d\Gamma(\vec{x}) \;+\; \int_\Omega p(\vec{x}) \left[\Delta \chi(\vec{x}) + k^2 \chi(\vec{x})\right] d\Omega(\vec{x}) \;=\; 0\,.\end{aligned} \tag{10}$$

Often, Eq. (9) represents the starting point for conventional finite element discretizations, e.g. Galerkin method. The second part (lower row) consists of a domain integral and a boundary integral. Similarly, the second part of Eq. (10) consists of one domain integral and two boundary integrals. This domain integral can be transformed into an integral-free term by using fundamental solutions $G(\vec{x}, \vec{y})$ in the sense of distributions. Function G represents the solution of the equation

$$\Delta G(\vec{x}, \vec{y}) + k^2 G(\vec{x}, \vec{y}) = \delta(\vec{x}, \vec{y}). \tag{11}$$

It is known as free-space Green's function as well, whereas $\delta(\vec{x}, \vec{y})$ is the Dirac or delta function at the origin $\vec{y}$. In terms of physics, $G(\vec{x}, \vec{y})$ can be understood as the sound pressure distribution according to a point source (monopole) in $\vec{y}$. Together with the harmonic time-dependence of $\mathrm{e}^{-\mathrm{i}\omega t}$, it represents an outgoing wave. We can write G as

$$
\begin{aligned}
G(\vec{x},\vec{y}) &= -\frac{1}{2k}\sin\left[kr(\vec{x},\vec{y})\right] && \vec{x},\vec{y}\in\mathbb{R}^1 \ , \\
G(\vec{x},\vec{y}) &= \frac{\mathrm{i}}{4}H_0^1\left(kr(\vec{x},\vec{y})\right) && \vec{x},\vec{y}\in\mathbb{R}^2 \ \text{and} \\
G(\vec{x},\vec{y}) &= \frac{1}{4\pi}\frac{\mathrm{e}^{\mathrm{i}kr(\vec{x},\vec{y})}}{r(\vec{x},\vec{y})} && \vec{x},\vec{y}\in\mathbb{R}^3 \ .
\end{aligned}
\tag{12}
$$

with r being the Euclidean distance between field point $\vec{x}$ and source point $\vec{y}$ as $r(\vec{x},\vec{y}) = |\vec{x}-\vec{y}|$. Note that the fundamental solutions are different when the time dependence is chosen to be $\mathrm{e}^{\mathrm{i}\omega t}$.

Applying the property of the fundamental solution and the delta function, we find

$$
\begin{aligned}
&\int_\Omega p(\vec{x})\left[\Delta G(\vec{x},\vec{y}) + k^2 G(\vec{x},\vec{y})\right] d\Omega(\vec{x}) = \\
&\qquad = \int_\Omega p(\vec{x})\left[\delta(\vec{x},\vec{y})\right] d\Omega(\vec{x}) = c(\vec{y})p(\vec{y}) \ .
\end{aligned}
\tag{13}
$$

The coefficient $c(\vec{y})$ is determined by the location of $\vec{y}$. It is

$$
\begin{aligned}
c(\vec{y}) &= 1 \quad \text{for} \quad \vec{y}\in\Omega \\
c(\vec{y}) &= 0 \quad \text{for} \quad \vec{y}\in\Omega_c \\
0 < c(\vec{y}) &< 1 \quad \text{for} \quad \vec{y}\in\Gamma
\end{aligned}
\tag{14}
$$

while the value of c is 0.5 if the $\vec{y}$ is a point on a smooth surface. It takes other values if $\vec{y}$ is located on an edge or a corner. With this, Eq. (10) is rewritten as

$$
c(\vec{y})p(\vec{y}) + \int_\Gamma \frac{\partial G(\vec{x},\vec{y})}{\partial n(\vec{x})} p(\vec{x}) d\Gamma(\vec{x}) = \int_\Gamma G(\vec{x},\vec{y})\, a v_f(\vec{x}) d\Gamma(\vec{x}). \tag{15}
$$

Equation (15) is known as representation formula since the sound pressure at an arbitrary point in Ω and Γ can be evaluated just with the knowledge of the boundary data. For $\vec{y}\in\Gamma$, it is known as the Kirchhoff–Helmholtz (boundary) integral equation. Note that plus and minus signs of either the first term or the second and the third term may be different if the direction of the normal vector is chosen in opposite direction. Similarly, these signs may change due to the choice of the harmonic time dependence and/or a negative sign on the right hand-side of Eq. (11), see for more details (Marburg 2016a). Therein, it was shown that a large fraction of the boundary element community is somehow confused about the correct choice of the signs. This will be discussed again later.

In Eq. (15), the integral on the right hand-side is also known as the single layer potential whereas the integral on the left hand-side is known as the double layer potential.

Before entering the discretization process, it will be useful to incorporate the boundary condition (4) into the weak formulation (15). Furthermore, we substitute for the constant a as

$$a \;=\; s\,k \qquad \text{with} \qquad s \;=\; \mathrm{i}\,\varrho_0\,c\;, \tag{16}$$

which explicitly shows wavenumber dependency. Then, Eq. (15) modifies to

$$\begin{aligned} c(\vec{\boldsymbol{y}})\,p(\vec{\boldsymbol{y}}) + \int_\Gamma \frac{\partial G(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}})}{\partial n(\vec{\boldsymbol{x}})}\,p(\vec{\boldsymbol{x}})\,d\Gamma(\vec{\boldsymbol{x}}) \;&=\\ =\; sk\int_\Gamma G(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}})\,\left[v_s(\vec{\boldsymbol{x}}) + Y(\vec{\boldsymbol{x}})\,p(\vec{\boldsymbol{x}})\right] d\Gamma(\vec{\boldsymbol{x}})\;. \end{aligned} \tag{17}$$

In case of $\vec{\boldsymbol{y}} \in \Gamma$ and, thus, $0 < c(\vec{\boldsymbol{y}}) < 1$, Eq. (17) represents a Fredholm integral equation of the second kind. This becomes more obvious if the part of the right hand-side integral which includes the sound pressure p is moved into the integral on the left hand-side

$$\begin{aligned} c(\vec{\boldsymbol{y}})\,p(\vec{\boldsymbol{y}}) + \int_\Gamma \left[\frac{\partial G(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}})}{\partial n(\vec{\boldsymbol{x}})} - sk\,G(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}})Y(\vec{\boldsymbol{x}})\right] p(\vec{\boldsymbol{x}})\,d\Gamma(\vec{\boldsymbol{x}}) \;&=\\ =\; sk\int_\Gamma G(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}})\,v_s(\vec{\boldsymbol{x}})\,d\Gamma(\vec{\boldsymbol{x}})\;. \end{aligned} \tag{18}$$

A boundary element formulation requires discretization of the Fredholm integral equation (18). For that, we introduce another test function $\tilde{\chi}(\vec{\boldsymbol{y}})$ such that

$$\begin{aligned} &\int_\Gamma \tilde{\chi}(\vec{\boldsymbol{y}})\,c(\vec{\boldsymbol{y}})\,p(\vec{\boldsymbol{y}})\,d\Gamma(\vec{\boldsymbol{y}}) \;+\\ &+\int_\Gamma \tilde{\chi}(\vec{\boldsymbol{y}})\left\{\int_\Gamma \left[\frac{\partial G(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}})}{\partial n(\vec{\boldsymbol{x}})} - sk\,G(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}})Y(\vec{\boldsymbol{x}})\right] p(\vec{\boldsymbol{x}})\,d\Gamma(\vec{\boldsymbol{x}})\right\} d\Gamma(\vec{\boldsymbol{y}}) =\\ &= sk\int_\Gamma \tilde{\chi}(\vec{\boldsymbol{y}})\left\{\int_\Gamma G(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}})\,v_s(\vec{\boldsymbol{x}})\,d\Gamma(\vec{\boldsymbol{x}})\right\} d\Gamma(\vec{\boldsymbol{y}})\,. \end{aligned} \tag{19}$$

Equation (19) is the basis for collocation and Galerkin discretization using boundary elements.

2.2 Discretization Process

Approximation: Independent of the discretization method, we formulate approximations of our physical quantities. First of all, we approximate the sound pressure $p(\vec{\boldsymbol{x}})$ as

$$p(\vec{\boldsymbol{x}}) \;=\; \sum_{l=1}^{N} \varphi_l(\vec{\boldsymbol{x}})\,p_l \;=\; \boldsymbol{\varphi}^T(\vec{\boldsymbol{x}})\,\boldsymbol{p}\;, \tag{20}$$

where p_l represents the discrete sound pressure at point $\vec{\boldsymbol{x}}_l$ and φ_l is the lth basis function for our approximation. Further, we assume that similar approximations are formulated for the particle velocity of the structure v_s and the boundary admittance Y

$$v_s(\vec{\boldsymbol{x}}) = \sum_{j=1}^{\bar{N}} \bar{\varphi}_j(\vec{\boldsymbol{x}})\, v_{s\,j} = \bar{\boldsymbol{\varphi}}^T(\vec{\boldsymbol{x}})\boldsymbol{v}_s \qquad \text{and}$$
$$Y(\vec{\boldsymbol{x}}) = \sum_{k=1}^{\tilde{N}} \tilde{\varphi}_k(\vec{\boldsymbol{x}})\, Y_k = \tilde{\boldsymbol{\varphi}}^T(\vec{\boldsymbol{x}})\boldsymbol{Y}\,. \tag{21}$$

If v_s and Y are explicitly known, these approximations are not necessary for evaluation of the boundary integrals in Eq. (18). However, there are many practical cases where the structural particle velocity is the result of a finite element simulation and available only as piecewise defined function. Similarly, the boundary admittance may vary locally or result from other evaluations which motivate the piecewise approximation.

The number of basis functions φ_l, $\bar{\varphi}_j$ and $\tilde{\varphi}_k$ is given by N, $\bar{N}$ and $\tilde{N}$, respectively. If the particle velocity of the structure and the boundary admittance are known functions, N accounts for degree of freedom. Herein, this coincides with the number of nodes of the finite or boundary element mesh. $\bar{N}$ and $\tilde{N}$ may be equal to each other. For BEM with discontinuous boundary elements, it is common that $\bar{N} = \tilde{N} = N$. In what follows, this will be assumed, meaning that there will not be distinction between the different basis functions φ, $\bar{\varphi}$ and $\tilde{\varphi}$.

In engineering literature, the discretization procedure of the integral equation, cf. Eq. (18), is often omitted or, at least, is abridged such that it can't be identified. Moreover, there are many engineering articles which prefer using the categories of direct and indirect approaches. Often, the direct approach is automatically associated with collocation whereas the indirect approach seems to be virtually linked to the Galerkin discretization. Herein, we limit our consideration to the direct approach. This, however, does not prohibit the use of either collocation or Galerkin discretization methods. Even other discretization methods can be used, e.g. Nyström methods and least squares methods. Similarly, the indirect approach allows different methods of discretization including collocation and Galerkin methods which are the most commonly used techniques for practical applications of the boundary element method.
Collocation: The collocation method requires substituting the Dirac function $\delta(\vec{\boldsymbol{y}}, \vec{\boldsymbol{z}})$ for the test function $\tilde{\chi}(\vec{\boldsymbol{y}})$ in Eq. (19). It modifies to

$$\int_\Gamma \delta(\vec{y},\vec{z})\, c(\vec{y})\, p(\vec{y})\, d\Gamma(\vec{y}) \; + \\ + \int_\Gamma \delta(\vec{y},\vec{z}) \left\{ \int_\Gamma \left[\frac{\partial G(\vec{x},\vec{y})}{\partial n(\vec{x})} \; - \; s\,k\, G(\vec{x},\vec{y})\, Y(\vec{x}) \right] p(\vec{x}) d\Gamma(\vec{x}) \right\} d\Gamma(\vec{y}) \;\; = \\ = \;\; s\,k \int_\Gamma \delta(\vec{y},\vec{z}) \left\{ \int_\Gamma G(\vec{x},\vec{y})\, v_s(\vec{x})\, d\Gamma(\vec{x}) \right\} d\Gamma(\vec{y})\,. \tag{22}$$

The outer integration is known analytically, cf. Eq. (13). It yields

$$c(\vec{z})p(\vec{z}) + \int_\Gamma \left[\frac{\partial G(\vec{x},\vec{z})}{\partial n(\vec{x})} \; - \; sk\, G(\vec{x},\vec{z}) Y(\vec{x}) \right] p(\vec{x}) d\Gamma(\vec{x}) \;\; = \\ = \;\; sk \int_\Gamma G(\vec{x},\vec{z})\, v_s(\vec{x}) d\Gamma(\vec{x})\,, \tag{23}$$

which is basically the same expression as shown in Eq. (18). The major difference between Eqs. (18) and (23) is that the former is actually a continuous integral equation whereas the latter is valid just for the discrete point $\vec{z}$. This means that the integral equation is fulfilled at a number of discrete points, i.e. collocation points $\vec{z}_l$. It is common practice that the collocation points coincide with the nodes of the piecewise formulated approximation of the sound pressure as shown in Eq. (20). For further considerations, we assume that $\varphi_l(\vec{z}_k) = \delta_{lk}$ where δ_{lk} is the Kronecker symbol with $\delta_{lk} = 0$ for $l \neq k$ and $\delta_{lk} = 1$ for $l = k$. Then, applying the approximation of Eqs. (20) and (21) yields

$$c(\vec{z}_l)\, p_l \; + \\ + \int_\Gamma \left\{ \frac{\partial G(\vec{x},\vec{z}_l)}{\partial n(\vec{x})} - skG(\vec{x},\vec{z}_l) \left[\sum_{j=1}^{N} \varphi_j(\vec{x}) Y_j \right] \right\} \left[\sum_{k=1}^{N} \varphi_k(\vec{x})\, p_k \right] d\Gamma(\vec{x}) = \\ = \;\; sk \int_\Gamma G(\vec{x},\vec{z}_l) \left[\sum_{m=1}^{N} \varphi_m(\vec{x}) v_{sm} \right] d\Gamma(\vec{x}). \tag{24}$$

To simplify this equation, we introduce matrices. Matrix $\boldsymbol{G}$ is the system matrix of the single layer potential as

$$g_{lj} \;\; = \;\; s\,k \int_\Gamma G(\vec{x},\vec{z}_l)\, \varphi_j(\vec{x})\, d\Gamma(\vec{x})\,, \tag{25}$$

matrix $\boldsymbol{H}$ contains the integral-free term and the contribution of the double layer potential as

$$
\begin{aligned}
h_{lj} &= c(\vec{z}_l)\,\delta_{lj} + \bar{h}_{lj} = \\
&= c(\vec{z}_l)\,\delta_{lj} + \int_\Gamma \frac{\partial G(\vec{x},\vec{z}_l)}{\partial n(\vec{x})}\,\varphi_j(\vec{x})\,d\Gamma(\vec{x})
\end{aligned}
\tag{26}
$$

and matrix $\boldsymbol{D}$ which contains the boundary admittance terms as

$$
\begin{aligned}
d_{lj} &= s\,k \int_\Gamma G(\vec{x},\vec{z}_l)\,\left[\boldsymbol{\varphi}^T(\vec{x})\,\boldsymbol{Y}\right]\,\varphi_j(\vec{x})\,d\Gamma(\vec{x}) = \\
&= \mathrm{i}\,k \int_\Gamma G(\vec{x},\vec{z}_l)\,\left[\boldsymbol{\varphi}^T(\vec{x})\,\tilde{\boldsymbol{Y}}\right]\,\varphi_j(\vec{x})\,d\Gamma(\vec{x})\,,
\end{aligned}
\tag{27}
$$

where the normalized admittance $\tilde{Y}$ is introduced. Equation (24) in matrix form is written as

$$
(\boldsymbol{H} - \boldsymbol{D})\;\boldsymbol{p} \;=\; \boldsymbol{G}\,\boldsymbol{v}_s \;=\; \boldsymbol{f}\,. \tag{28}
$$

The system matrices $\boldsymbol{G}$, $\boldsymbol{H}$ and $\boldsymbol{D}$ are neither Hermitian nor positive definite in general. There are examples in literature where $\boldsymbol{D} = \boldsymbol{G}\boldsymbol{Y}$ with the diagonal matrix $\boldsymbol{Y}$. This form requires some specified conditions as for example piecewise constant approximation of the boundary admittance and, more general, discontinuous approximation of the sound pressure. In upcoming sections, the version with the term $\boldsymbol{GY}$ will usually be used.

It should be noted that although three matrices are used in the formulation, it is possible (and not very difficult to be organized in a computer code) to set up all system matrices at once and keep only one system matrix in memory. Setup of the system matrices requires $O(N^2)$ floating point operations. With a complete system matrix in memory, the memory requirements are also $O(N^2)$.

Galerkin Method: The classical Galerkin method requires use of the basis functions φ_l for approximation of the sound pressure and for the test function $\tilde{\chi}(\vec{y})$ in Eq. (19) as

$$
\begin{aligned}
&\int_\Gamma \varphi_l(\vec{y})\,c(\vec{y})\,p(\vec{y})\,d\Gamma(\vec{y}) + \\
&\quad + \int_\Gamma \varphi_l(\vec{y}) \left\{ \int_\Gamma \left[\frac{\partial G(\vec{x},\vec{y})}{\partial n(\vec{x})} - s\,k\,G(\vec{x},\vec{y})\,Y(\vec{x}) \right] p(\vec{x}) d\Gamma(\vec{x}) \right\} d\Gamma(\vec{y}) = \\
&\qquad = s\,k \int_\Gamma \varphi_l(\vec{y}) \left\{ \int_\Gamma G(\vec{x},\vec{y})\,v_s(\vec{x})\,d\Gamma(\vec{x}) \right\} d\Gamma(\vec{y})\,.
\end{aligned}
\tag{29}
$$

This time, the double surface integral does not vanish as it did for collocation. Simplification of the first term becomes possible since the integration over piecewise smooth surface elements allows to set $c(\vec{y}) = 1/2$. Applying the approximation of Eqs. (20) and (21) gives (omitting dependencies on $\vec{x}$ and $\vec{y}$)

$$\frac{1}{2}\int_\Gamma \varphi_l \left[\sum_{j=1}^{N} \varphi_j\, p_j\right] d\Gamma + \\ + \int_\Gamma \varphi_l \left\langle \int_\Gamma \left\{ \frac{\partial G}{\partial n_x} - s\,k\,G\left[\sum_{k=1}^{N} \varphi_k\, Y_k\right]\right\}\left[\sum_{j=1}^{N} \varphi_j\, p_j\right] d\Gamma_x \right\rangle d\Gamma \;= \\ = \; s\,k \int_\Gamma \varphi_l \left\{ \int_\Gamma G \left[\sum_{j=1}^{N} \varphi_j\, v_{s\,j}\right] d\Gamma \right\} d\Gamma\,. \tag{30}$$

Similar to the collocation method, we introduce matrices $\boldsymbol{G}$, $\boldsymbol{H}$, *and* $\boldsymbol{D}$. We write the system matrix of the single layer potential $\boldsymbol{G}$ as

$$g_{lj} \;=\; s\,k \int_\Gamma \int_\Gamma G(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}})\, \varphi_l(\vec{\boldsymbol{y}})\, \varphi_j(\vec{\boldsymbol{x}})\, d\Gamma(\vec{\boldsymbol{x}}) d\Gamma(\vec{\boldsymbol{y}})\,, \tag{31}$$

the matrix of the double layer potential $\boldsymbol{H}$ as

$$h_{lj} = \frac{1}{2}\int_\Gamma \varphi_l(\vec{\boldsymbol{y}})\, \varphi_j(\vec{\boldsymbol{y}}) d\Gamma(\vec{\boldsymbol{y}}) + \int_\Gamma \int_\Gamma \frac{\partial G(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}})}{\partial n(\vec{\boldsymbol{x}})}\, \varphi_l(\vec{\boldsymbol{y}})\, \varphi_j(\vec{\boldsymbol{x}})\, d\Gamma(\vec{\boldsymbol{x}}) d\Gamma(\vec{\boldsymbol{y}}) \\ = \frac{1}{2}\theta_{lj} + \int_\Gamma \int_\Gamma \frac{\partial G(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}})}{\partial n(\vec{\boldsymbol{x}})}\, \varphi_l(\vec{\boldsymbol{y}})\, \varphi_j(\vec{\boldsymbol{x}})\, d\Gamma(\vec{\boldsymbol{x}}) d\Gamma(\vec{\boldsymbol{y}}), \tag{32}$$

the matrix which contains the boundary admittance terms $\boldsymbol{D}$ as

$$d_{lj} = s\,k \int_\Gamma \int_\Gamma G(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}}) \left[\boldsymbol{\varphi}^T(\vec{\boldsymbol{x}})\, \boldsymbol{Y}\right] \varphi_l(\vec{\boldsymbol{y}}) \varphi_j(\vec{\boldsymbol{x}})\, d\Gamma(\vec{\boldsymbol{x}}) d\Gamma(\vec{\boldsymbol{y}}) \;= \\ = \mathrm{i}\,k \int_\Gamma \int_\Gamma G(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}}) \left[\boldsymbol{\varphi}^T(\vec{\boldsymbol{x}})\, \tilde{\boldsymbol{Y}}\right] \varphi_l(\vec{\boldsymbol{y}}) \varphi_j(\vec{\boldsymbol{x}})\, d\Gamma(\vec{\boldsymbol{x}}) d\Gamma(\vec{\boldsymbol{y}})\,. \tag{33}$$

The normalized boundary admittance $\tilde{Y}$ has been used again. In Eq. (32), the boundary mass matrix $\boldsymbol{\Theta}$ has been introduced such that

$$\theta_{lj} \;=\; \int_\Gamma \varphi_l(\vec{\boldsymbol{y}})\, \varphi_j(\vec{\boldsymbol{y}}) d\Gamma(\vec{\boldsymbol{y}})\,. \tag{34}$$

There is no physical relevance as a mass matrix for $\boldsymbol{\Theta}$. The term is mainly based on the definition of the mass matrix in finite element formulations. The matrix $\boldsymbol{\Theta}$ is of a practical relevance at many occasions. Finally, we write the system of equations in matrix form as

$$(\boldsymbol{H} - \boldsymbol{D})\; \boldsymbol{p} \;=\; \boldsymbol{G}\, \boldsymbol{v}_s \;=\; \boldsymbol{f}\,. \tag{35}$$

Similar to collocation BEM, the system matrices $\boldsymbol{G}$, $\boldsymbol{H}$ and $\boldsymbol{D}$ are neither Hermitian nor positive definite in general. It is possible to set up a Hermitian system of equations

for the Galerkin BEM. Examples for this have been shown in Chen et al. (2008) and in Gaul et al. (2008).

2.3 Solution of the Linear System of Equations

Formally, the systems of Eqs. (28) and (35) are solved by inversion of the system matrix as

$$\boldsymbol{p} \;=\; (\boldsymbol{H} \,-\, \boldsymbol{D})^{-1}\, \boldsymbol{G}\, \boldsymbol{v}_s \;=\; (\boldsymbol{H} \,-\, \boldsymbol{D})^{-1}\, \boldsymbol{f}\,. \tag{36}$$

However, the inverse matrix is usually not evaluated since this is computationally too costly if the degree of freedom N is getting large. Usually, the degree of freedom is the same as the number of nodes. The systems of Eqs. (28) and (35) are solved either by using a direct solver or an iterative solver.

Direct solvers, such as Gaussian elimination which have a complexity of $O(N^3)$ floating point operations can be suitable for smaller problems of up to a few thousand degrees of freedom.

For larger models, it is recommended to use iterative solvers. There is a variety of Krylov subspace methods which are suited for non-Hermitian systems of equations. Some of these methods have been tested for boundary element techniques in acoustics and compared to each other, cf. Marburg and Schneider (2003a), Chen et al. (2000) and Sakuma et al. (2008). Although not exclusively used, the most popular iterative solver seems to be the Generalized Minimal Residual (GMRes) technique which was proposed by Saad and Schultz (1986).

Iterative solvers require evaluation of at least one matrix-vector product in each step. Since the matrix-vector product requires $O(N^2)$ floating point operations, the entire solution of the system of equations possesses this complexity. When keeping the entire system matrix in memory, the memory requirements are $O(N^2)$ and, thus, similar to the number of operations.

The mathematical and the engineering literature knows many suitable preconditioners. The author has had good experiences with incomplete LU decomposition, i.e. with an iLU preconditioner, cf. Schneider and Marburg (2003), Chen et al. (2000), see also the discussion in Sakuma et al. (2008).

2.4 One-Dimensional Example

Boundary value problem and common solution technique: In a one-dimensional example, the partial differential equation (3) changes into an ordinary differential equation as

$$\frac{d^2 p(x)}{dx^2} + k^2 p(x) = 0 \qquad \text{with} \qquad x \in [0, l]\,. \tag{37}$$

The boundary conditions are formulated for the points at $x = 0$ and $x = l$

$$\begin{aligned} \frac{dp(x=0)}{dx} &= -sk(v_0 + Y_0 p(x=0)) \qquad \text{and} \\ \frac{dp(x=l)}{dx} &= sk(v_l + Y_l p(x=l)) \end{aligned} \tag{38}$$

where the negative sign is introduced at $x = 0$ since the outward normal is pointing into negative direction. The particle velocities and the admittance values at both ends are denoted by v_0, v_l, Y_0 and Y_l, respectively. The solution of (37) is known to be

$$p(x) = A\mathrm{e}^{\mathrm{i}kx} + B\mathrm{e}^{-\mathrm{i}kx} \tag{39}$$

where the constants A and B are adjusted such that the boundary conditions are fulfilled

$$\begin{aligned} A &= -\rho c \frac{v_0(1+\tilde{Y}_l)e^{-\mathrm{i}kl} + v_l(1-\tilde{Y}_0)}{(1+\tilde{Y}_0+\tilde{Y}_l+\tilde{Y}_0\tilde{Y}_l)e^{-\mathrm{i}kl} - (1-\tilde{Y}_0-\tilde{Y}_l+\tilde{Y}_0\tilde{Y}_l)e^{\mathrm{i}kl}} \\ B &= -\rho c \frac{v_0(1-\tilde{Y}_l)e^{\mathrm{i}kl} + v_l(1+\tilde{Y}_0)}{(1+\tilde{Y}_0+\tilde{Y}_l+\tilde{Y}_0\tilde{Y}_l)e^{-\mathrm{i}kl} - (1-\tilde{Y}_0-\tilde{Y}_l+\tilde{Y}_0\tilde{Y}_l)e^{\mathrm{i}kl}} \end{aligned} \tag{40}$$

where the normalized boundary admittance $\tilde{Y} = \rho_0 c Y$ is used again. A detailed study of the eigenvalue problem for this case has been carried out by the author in Marburg (2005) where it is used for as comparison with finite element solutions.

Boundary data solution using Green's function: A suitable Green's function or fundamental solution for the 1d problem has been given in Eq. (12). Note that this solution is not unique as it is usually the case for particular integrals. They are just required to solve the inhomogeneous differential equation. It is useful for our purposes to rewrite Eq. (12) in a slightly different form

$$G(x, y) = -\frac{1}{2k} \sin(k|x-y|). \tag{41}$$

Different from the Green's functions in 2d and 3d, the 1d solution does not expose a singularity for $x = y$. However, it exposes a kink for $x = y$ and is therefore not uniquely differentiable at this point. For evaluation of the normal derivative it is useful to determine the derivative with respect to x as

$$\frac{dG(x, y)}{dx} = -\frac{1}{2} \cos(k|x-y|) \frac{d|x-y|}{dx}. \tag{42}$$

It is necessary to distinguish two cases

$$\frac{dG(x,y)}{dx} = \frac{1}{2}\cos\left(k\left(x-y\right)\right) \quad \text{for} \quad x < y$$
$$\frac{dG(x,y)}{dx} = -\frac{1}{2}\cos\left(k\left(x-y\right)\right) \quad \text{for} \quad x > y. \tag{43}$$

There are different ways to formulate a boundary equation: One is based on partial integration of a weighted residual function of the 1d Helmholtz equation in a similar way as shown in a more general way in Eqs. (8)–(10). Another one is an adjustment of Eq. (18) to the 1d case. In 1d, the boundary integrals vanish since the boundary consists of two discrete points only. Thus, the integral equation becomes

$$\begin{aligned} p(y) - \frac{dG(0,y)}{dx}p_0 + \frac{dG(l,y)}{dx}p_l - G(0,y)skY_0p_0 - G(l,y)skY_lp_l \\ = G(0,y)skv_0 + G(l,y)skv_l. \end{aligned} \tag{44}$$

The system of equations is set up by writing one equation for $y = 0$ and another one for $y = l$ which are the collocation points now. This yields

$$\begin{aligned} \left[1 - \frac{dG(x \to -0, 0)}{dx} - G(0,0)skY_0\right]p_0 + \left[\frac{dG(l,0)}{dx} - G(l,0)skY_l\right]p_l \\ = G(0,0)skv_0 + G(l,0)skv_l \end{aligned} \tag{45}$$

and

$$\begin{aligned} \left[-\frac{dG(0,l)}{dx} - G(0,l)skY_0\right]p_0 + \left[1 + \frac{dG(x \to +0+l, l)}{dx} - G(l,l)skY_l\right]p_l \\ = G(0,l)skv_0 + G(l,l)skv_l \end{aligned} \tag{46}$$

where the derivative of the Green's function at $x = y$ is evaluated from the outer side since the normal vector is pointing out of the domain Ω. It is easily possible to evaluate all coefficients of these two equations. For the Green's function itself, we get

$$\begin{aligned} G(0,0) &= -\frac{1}{2k}\sin\left(k\left|0-0\right|\right) = 0 \\ G(0,l) &= -\frac{1}{2k}\sin\left(k\left|0-l\right|\right) = -\frac{1}{2k}\sin\left(kl\right) \\ G(l,0) &= -\frac{1}{2k}\sin\left(k\left|l-0\right|\right) = -\frac{1}{2k}\sin\left(kl\right) \\ G(l,l) &= -\frac{1}{2k}\sin\left(k\left|l-l\right|\right) = 0 \end{aligned} \tag{47}$$

and for the derivatives

$$\begin{aligned}
\frac{dG(x \to -0,0)}{dx} &= \frac{1}{2}\cos\left(k\left(0-0\right)\right) = \frac{1}{2} \\
\frac{dG(x=0,l)}{dx} &= \frac{1}{2}\cos\left(k\left(0-l\right)\right) = \frac{1}{2}\cos\left(kl\right) \\
\frac{dG(x=l,0)}{dx} &= -\frac{1}{2}\cos\left(k\left(l-0\right)\right) = -\frac{1}{2}\cos\left(kl\right) \\
\frac{dG(x \to -0+l,l)}{dx} &= -\frac{1}{2}\cos\left(k\left(l-l\right)\right) = -\frac{1}{2}.
\end{aligned} \tag{48}$$

Applying these results to Eqs. (45) and (46) yields

$$\begin{aligned}
\frac{1}{2}p_0 - \left[\frac{1}{2}\cos(kl) - \frac{1}{2}\sin(kl)sY_l\right]p_l &= -\frac{1}{2}\sin(kl)sv_l \\
\frac{1}{2}p_l - \left[\frac{1}{2}\cos(kl) - \frac{1}{2}\sin(kl)sY_0\right]p_0 &= -\frac{1}{2}\sin(kl)sv_0.
\end{aligned} \tag{49}$$

Now it is possible to write the system of equations similar to (28) in matrix form as

$$\left(\boldsymbol{H} - \boldsymbol{GY}\right)\boldsymbol{p} = \boldsymbol{Gv}_s \tag{50}$$

where

$$\begin{aligned}
\boldsymbol{G} &= s\begin{bmatrix} 0 & -\sin(kl) \\ -\sin(kl) & 0 \end{bmatrix}, \quad \boldsymbol{H} = \begin{bmatrix} 1 & -\cos(kl) \\ -\cos(kl) & 1 \end{bmatrix}, \\
\boldsymbol{Y} &= \begin{bmatrix} Y_0 & 0 \\ 0 & Y_l \end{bmatrix}, \quad \boldsymbol{p} = \begin{bmatrix} p_0 \\ p_l \end{bmatrix} \text{ and } \boldsymbol{v}_s = \begin{bmatrix} v_0 \\ v_l \end{bmatrix}.
\end{aligned} \tag{51}$$

Finally, the 1d problem allowed to define all the matrices as formulated in the more general multi-dimensional formulation. Solution requires inversion of the system matrix $\boldsymbol{H} - \boldsymbol{GY}$

$$\left(\boldsymbol{H} - \boldsymbol{GY}\right)^{-1} = \frac{1}{N}\begin{bmatrix} 1 & \cos(kl) - \mathrm{i}\tilde{Y}_l\sin(kl) \\ \cos(kl) - \mathrm{i}\tilde{Y}_0\sin(kl) & 1 \end{bmatrix} \tag{52}$$

with

$$N = \sin^2(kl) + \mathrm{i}(\tilde{Y}_0 + \tilde{Y}_l)\sin(kl)\cos(kl) + \tilde{Y}_0\tilde{Y}_l\sin^2(kl). \tag{53}$$

This is leading to the sound pressure at the two boundary points

$$\boldsymbol{p} = \begin{bmatrix} p_0 \\ p_l \end{bmatrix} = -s \begin{bmatrix} \dfrac{\left[\cos(kl) - \mathrm{i}\tilde{Y}_l \sin(kl)\right] v_0 + v_l}{\sin(kl) + \mathrm{i}(\tilde{Y}_0 + \tilde{Y}_l)\cos(kl) + \tilde{Y}_0\tilde{Y}_l \sin(kl)} \\ \dfrac{v_0 + \left[\cos(kl) - \mathrm{i}\tilde{Y}_0 \sin(kl)\right] v_l}{\sin(kl) + \mathrm{i}(\tilde{Y}_0 + \tilde{Y}_l)\cos(kl) + \tilde{Y}_0\tilde{Y}_l \sin(kl)} \end{bmatrix} . \tag{54}$$

Usually, this solution accounts for the solution of the boundary value problem. Further computations such as field point evaluations are just postprocessing. The reader is invited to confirm that the solutions for $p(y = 0)$ and $p(y = l)$ are identical in Eqs. (39) and (54). The author has tested this numerically.

3 Sources and Incident Wave Fields

3.1 A General Approach

The previous considerations have started from the homogeneous Helmholtz equation with inhomogeneous boundary conditions. Now, we consider problems with sources, e.g. monopole sources and, further, incident waves.

Assume the source q located somewhere in Ω, then the Helmholtz equation (3) becomes inhomogeneous as

$$\Delta p(\vec{\boldsymbol{x}}) + k^2 p(\vec{\boldsymbol{x}}) = q(\vec{\boldsymbol{x}}) . \tag{55}$$

Since Eq. (55) is a linear inhomogeneous partial differential equation, its solution can be constructed by superimposing the solution of the homogeneous Helmholtz equation (3) which is also known as complementary solution p^c and a particular solution p^p which solves the inhomogeneous Eq. (55). In acoustics, the complementary solution is usually referred to as the scattered pressure p^s while the particular solution is usually referred to as the incident pressure field p^i. Denoting the entire sound pressure by p, it can be represented as

$$p(\vec{\boldsymbol{x}}) = p^c(\vec{\boldsymbol{x}}) + p^p(\vec{\boldsymbol{x}}) = p^s(\vec{\boldsymbol{x}}) + p^i(\vec{\boldsymbol{x}}) \tag{56}$$

for the sound pressure and

$$v_f(\vec{\boldsymbol{x}}) = v_f^c(\vec{\boldsymbol{x}}) + v_f^p(\vec{\boldsymbol{x}}) = v_f^s(\vec{\boldsymbol{x}}) + v_f^i(\vec{\boldsymbol{x}}) \tag{57}$$

for the fluid particle velocity. The two most commonly applied sources are the monopole source and the incident wave field.

For the monopole source, we can formulate the source term in Eq. (55) as

$$q(\vec{\boldsymbol{x}}) = q_m(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}}) = C\,\delta(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}}) , \tag{58}$$

thus, looking very similar to the right hand side of Eq. (11). Actually, the Dirac function in (11) accounts for the inhomogeneity which leads to the fundamental solution or Green's function G. Similarly, the source term in Eq. (58) demands a particular solution as

$$p^i(\vec{x},\vec{y}) \;=\; \frac{C}{4\pi}\,\frac{\mathrm{e}^{\mathrm{i}kr(\vec{x},\vec{y})}}{r(\vec{x},\vec{y})} \qquad\qquad \vec{x},\vec{y}\in\mathbb{R}^3\;. \tag{59}$$

Assuming the monopole to be a pulsating sphere (radius R) with constant surface sound pressure p_0, the particular solution p^i can be written in terms of the radius $r \geq R$ as

$$p^i(r) \;=\; p_0\,\frac{R}{r}\,\mathrm{e}^{\mathrm{i}k(r-R)}\;, \tag{60}$$

whereas the monopole solution for a prescribed surface particle velocity v_{f_0} provides the sound pressure distribution as

$$p^i(r) \;=\; \varrho_0 c v_{f_0}\,\frac{R}{r}\,\frac{\mathrm{i}kR}{1-\mathrm{i}kR}\,\mathrm{e}^{\mathrm{i}k(r-R)}\;. \tag{61}$$

The particular solution of the particle velocity v_f^i for the monopole with surface pressure prescribed is

$$v_f^i(r) \;=\; \frac{\mathrm{i}p_0R}{\varrho_0\omega}\,\frac{1-\mathrm{i}kr}{r^2}\,\mathrm{e}^{\,\mathrm{i}k(r-R)}\,\frac{\partial r}{\partial n}\;, \tag{62}$$

whereas for the monopole with surface particle velocity prescribed, we get

$$v_f^i(r) \;=\; v_{f_0}\,\frac{1-\mathrm{i}kr}{1-\mathrm{i}kR}\,\frac{R^2}{r^2}\,\mathrm{e}^{\,\mathrm{i}k(r-R)}\,\frac{\partial r}{\partial n}\;. \tag{63}$$

Note that $\partial p^i(r=R)/\partial r = -a v_{f_0}$, where the negative sign results from the definition of the normal vector for v_f according to Eq. (5) and Fig. 1.

An analogous approach is possible for any other source or even any sink with either known right hand side q of Eq. (55) or known particular solution p^i. It might even be possible to experimentally determine the particular solution by measuring and analytically reconstructing the sound pressure distribution around an arbitrary radiator under free-field conditions. This can be an option for a loudspeaker in a frequency range where the monopole characteristics cannot be guaranteed.

Although the situation is different for a uniform incident plane wave field, it is treated in the same way as the monopole source above. The infinite uniform incident plane wave can be described by the particular solution

$$p^i(\vec{x}) \;=\; p_0\,\mathrm{e}^{\,\mathrm{i}(\vec{k}\cdot\vec{x}+\varphi_0)}\;. \tag{64}$$

The vector $\vec{\boldsymbol{k}}$ contains the components which determine the direction of the traveling waves. The magnitude of $\vec{\boldsymbol{k}}$ is the actual wavenumber k, hence $k = |\vec{\boldsymbol{k}}|$. The reference phase angle φ_0 should be suitably chosen. For the particle velocity this yields

$$v_f^i(\vec{\boldsymbol{x}}) \;=\; \frac{p_0}{\varrho_0 \omega}\, \vec{\boldsymbol{k}} \cdot \vec{\boldsymbol{n}}\, \mathrm{e}^{\,\mathrm{i}(\vec{\boldsymbol{k}}\cdot\vec{\boldsymbol{x}}+\varphi_0)} \;=\; \frac{p_0}{\varrho_0 c}\, \frac{\partial k}{\partial n}\, \mathrm{e}^{\mathrm{i}(\vec{\boldsymbol{k}}\cdot\vec{\boldsymbol{x}}+\varphi_0)}\,. \tag{65}$$

It can be easily seen that the particular solution in Eq. (64) fulfills the homogeneous Helmholtz equation. Thus, the right hand side q in Eq. (55) vanishes, $q = 0$. Consequently, the particular solution is just an additional part of the complementary solution which does not necessarily fulfill other conditions such as the Sommerfeld radiation condition (7). From the physical point of view, the infinite uniform incident wave field does not make sense since it assumes arbitrarily many (i.e. an infinite number of) energy sources and sinks.

Superposition of complementary solutions p^s, v_f^s and particular solutions p^i, v_f^i provides us with the complete solutions p, v_f. Substitutions can be made at any point of the finite and boundary element approaches in the previous sections which are valid for the complementary solutions only. It is reasonable to start this substitution at Eq. (15). Performing this, the boundary element systems of Eqs. (28) and (35) become

$$(\boldsymbol{H} \;-\; \boldsymbol{D})\, \boldsymbol{p} \;=\; \boldsymbol{G}\left(\boldsymbol{v}_s \;-\; \boldsymbol{v}_f^i\right) \;+\; \boldsymbol{H}\boldsymbol{p}^i \;=\; \boldsymbol{f} \;+\; \boldsymbol{f}^i\,. \tag{66}$$

It is noteworthy that the substitution of complete and particular solution by the complementary solution is required prior to application of the boundary condition since the latter is only valid for the actual physical magnitude. Equation (66) shows that the particular solution contributes to the right hand side only.

3.2 A Different Approach

The concept of source consideration of the previous subsection is valid for any formulation which is derived from the Helmholtz equation including finite element approaches. Formally, the inhomogeneous Helmholtz equation can be written in a weak formulation, integrated by parts twice and, thus, changing Eq. (15) into

$$\begin{aligned} c(\vec{\boldsymbol{y}})\, p(\vec{\boldsymbol{y}}) \;+\; \int_\Gamma \frac{\partial G(\vec{\boldsymbol{x}},\, \vec{\boldsymbol{y}})}{\partial n(\vec{\boldsymbol{x}})}\, p(\vec{\boldsymbol{x}}) d\Gamma(\vec{\boldsymbol{x}}) \;&=\; \\ =\quad \int_\Gamma G(\vec{\boldsymbol{x}},\, \vec{\boldsymbol{y}})\, a v_f(\vec{\boldsymbol{x}}) d\Gamma(\vec{\boldsymbol{x}}) \;&+\; \int_Q G(\vec{\boldsymbol{x}},\, \vec{\boldsymbol{y}})\, q(\vec{\boldsymbol{x}},\, \vec{\boldsymbol{z}}) d\,Q(\vec{\boldsymbol{x}})\,, \end{aligned} \tag{67}$$

where the free field Q is the combination of Ω and Ω_c, hence $\Omega \cup \Omega_c \cup Q$. Often, Eq. (67) is not suited for practical use.

It is useful to return to Eq. (15) which has introduced the Kirchhoff–Helmholtz integral equation. It is valid for the source free acoustic field. Therefore, we write

$$c(\vec{y})\,p^s(\vec{y}) \;+\; \int_\Gamma \frac{\partial G(\vec{x},\,\vec{y})}{\partial n(\vec{x})}\; p^s(\vec{x}) d\Gamma(\vec{x}) \;=\; \int_\Gamma G(\vec{x},\,\vec{y})\, a v_f^s(\vec{x}) d\Gamma(\vec{x})\,. \tag{68}$$

Incorporating (56) and (57) into (68) yields

$$\begin{aligned} c(\vec{y})\left[p(\vec{y}) - p^i(\vec{y})\right] \;+\; \int_\Gamma \frac{\partial G(\vec{x},\,\vec{y})}{\partial n(\vec{x})}\; \left[p(\vec{x}) - p^i(\vec{x})\right] d\Gamma(\vec{x}) \;=\; \\ =\; \int_\Gamma G(\vec{x},\,\vec{y})\, a \left[v_f(\vec{x}) - v_f^i(\vec{x})\right] d\Gamma(\vec{x}) \end{aligned} \tag{69}$$

which obviously leads to the system of equations shown in Eq. (66).

Alternatively, we may assume that the particular solution either fulfills the homogeneous Helmholtz equation, e.g. plane wave solution (64) and (65), or the source is located in Ω, e.g. the monopole solution in Eqs. (59)–(63) with $\vec{y} \in \Omega$. Then, the particular solution fulfills the homogeneous Helmholtz equation, and thus, the Kirchhoff–Helmholtz integral equation in the complementary domain Ω_c

$$\begin{aligned} -\,\tilde{c}(\vec{y})\,p^i(\vec{y}) \;+\; \int_\Gamma \frac{\partial G(\vec{x},\,\vec{y})}{\partial n(\vec{x})}\; p^i(\vec{x}) d\Gamma(\vec{x}) \;=\; \int_\Gamma G(\vec{x},\,\vec{y})\, a v_f^i(\vec{x}) d\Gamma(\vec{x}) \\ \text{for}\;\; \vec{y} \in \Omega_c\,. \end{aligned} \tag{70}$$

Note that the normal vector is pointing into the complementary domain. Further, notice $c(\vec{y})$ in Eq. (68) and $\tilde{c}(\vec{y})$ in Eq. (70) are related such that $c(\vec{y}) + \tilde{c}(\vec{y}) = 1$. This is obvious if either $\vec{y} \in \Omega$ or $\vec{y} \in \Omega_c$ but also holds for $\vec{y} \in \Gamma$, since c is the value for $\vec{y} \in \Omega$ approaching Γ and $\tilde{c}$ is the value for $\vec{y} \in \Omega_c$ approaching Γ. Summation of Eqs. (68) and (70) together with substituting $p - p^i$ for the complementary solution leads to

$$\begin{aligned} c(\vec{y})\,p(\vec{y}) + \int_\Gamma \frac{\partial G(\vec{x},\,\vec{y})}{\partial n(\vec{x})}\; p(\vec{x}) d\Gamma(\vec{x}) \;=\; \int_\Gamma G(\vec{x},\,\vec{y})\, a v_f(\vec{x}) d\Gamma(\vec{x}) \;+\; p^i(\vec{y}) \\ \text{for}\;\; \vec{y} \in \Gamma\,. \end{aligned} \tag{71}$$

Note that Eq. (71) can be derived from Eq. (67) for the case of an arbitrary three-dimensional monopole source, see also Eqs. (58) and (59). This yields

$$\begin{aligned} \int_Q G(\vec{x},\,\vec{y})\, q(\vec{x},\,\vec{z}) d Q(\vec{x}) = \int_Q G(\vec{x},\,\vec{y})\, C\, \delta(\vec{x},\,\vec{z}) d Q(\vec{x}) \;= \\ = C\, G(\vec{x},\,\vec{z}) \;=\; \frac{C}{4\pi}\, \frac{\mathrm{e}^{\,\mathrm{i}kr(\vec{x},\vec{z})}}{r(\vec{x},\,\vec{z})} \;=\; p^i(\vec{x},\,\vec{z})\,. \end{aligned} \tag{72}$$

Equation (71) is probably the most popular variant to consider sources in boundary element techniques. The final system of equations for the boundary element collocation method becomes

$$(\boldsymbol{H} - \boldsymbol{D})\,\boldsymbol{p} = \boldsymbol{G}\boldsymbol{v}_s + \boldsymbol{p}^i = \boldsymbol{f} + \boldsymbol{f}^i\,. \tag{73}$$

Again, the particular solution appears as an additional term on the right hand-side only. Interestingly, Eqs. (66) and (73) are fully equivalent to each other.

3.3 Source Distribution as Boundary Condition

Another very common practice is to apply the particular solution as a boundary condition for the problem. This can be easily explained for the system of Eq. (66). The complementary solution of the sound pressure $\boldsymbol{p}^s$ is substituted for $\boldsymbol{p} - \boldsymbol{p}^i$. This leads to

$$(\boldsymbol{H} - \boldsymbol{D})\,\boldsymbol{p}^s = \boldsymbol{G}\left(\boldsymbol{v}_s - \boldsymbol{v}_f^i\right) + \boldsymbol{D}\boldsymbol{p}^i = \boldsymbol{f} + \boldsymbol{f}^i. \tag{74}$$

Note the simplification in case of acoustically rigid surfaces where the matrix $\boldsymbol{D}$ vanishes. For the single scattering problem, i.e. $\vec{\boldsymbol{v}}_s = \boldsymbol{0}$, the negative particular solution of the particle velocity remains the only boundary condition. It can be seen as if the solution $\boldsymbol{p}^s$ must be chosen such that the boundary condition which is induced by the particular solution is balanced.

4 Sound Field Evaluation

4.1 Field Point Evaluation

So far, solution of the systems of equations in (28), (35), (66), (73), and (74) has only returned boundary data, i.e. the sound pressure at the boundary, while the particle velocity and the admittance data have been known beforehand since they are the boundary conditions. With the knowledge of the whole set of boundary data, it is possible to evaluate the sound pressure at every field point $\vec{\boldsymbol{y}} \in \Omega$ by adjusting the representation formula (15) as

$$p(\vec{\boldsymbol{y}}) = \int_\Gamma G(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}})\, sk v_f(\vec{\boldsymbol{x}}) d\Gamma(\vec{\boldsymbol{x}}) - \int_\Gamma \frac{\partial G(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}})}{\partial n(\vec{\boldsymbol{x}})}\, p(\vec{\boldsymbol{x}}) d\Gamma(\vec{\boldsymbol{x}}). \tag{75}$$

Based on the derivation in the previous sections, it is easy to accommodate sources and admittance boundary conditions. A discretized version of the representation formula can be written as

$$p(\vec{\boldsymbol{y}}) \;=\; \boldsymbol{g}^T(\vec{\boldsymbol{y}})\boldsymbol{v}_f - \boldsymbol{h}^T(\vec{\boldsymbol{y}})\boldsymbol{p} \tag{76}$$

where $\boldsymbol{v}_f$ and $\boldsymbol{p}$ account for the column matrices of the nodal data of fluid particle velocity and sound pressure, respectively. Column matrices $\boldsymbol{g}$ and $\boldsymbol{h}$ are very similarly evaluated as rows of the system matrices $\boldsymbol{G}$ and $\boldsymbol{H}$ with collocation

$$g_j(\vec{\boldsymbol{y}}) \;=\; s\,k \int_\Gamma G(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}})\, \varphi_j(\vec{\boldsymbol{x}})\, d\Gamma(\vec{\boldsymbol{x}}) \;, \tag{77}$$

and

$$h_j(\vec{\boldsymbol{y}}) \;=\; \int_\Gamma \frac{\partial G(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}})}{\partial n(\vec{\boldsymbol{x}})}\, \varphi_j(\vec{\boldsymbol{x}})\, d\Gamma(\vec{\boldsymbol{x}}). \tag{78}$$

Incorporation of admittance boundary conditions modifies (76) into

$$p(\vec{\boldsymbol{y}}) \;=\; \boldsymbol{g}^T(\vec{\boldsymbol{y}})\boldsymbol{v}_s - \left[\boldsymbol{h}^T(\vec{\boldsymbol{y}}) - \boldsymbol{d}^T(\vec{\boldsymbol{y}})\right]\boldsymbol{p} \tag{79}$$

with

$$d_j(\vec{\boldsymbol{y}}) = s\,k \int_\Gamma G(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}})\, \left[\boldsymbol{\varphi}^T(\vec{\boldsymbol{x}})\boldsymbol{Y}\right]\, \varphi_j(\vec{\boldsymbol{x}})\, d\Gamma(\vec{\boldsymbol{x}}). \tag{80}$$

Similar to what was mentioned in the context of Eq. (28), $\boldsymbol{d}$ is often written as $\boldsymbol{Y}\boldsymbol{g}$. This will be used in what follows. Thus,

$$p(\vec{\boldsymbol{y}}) \;=\; \boldsymbol{g}^T(\vec{\boldsymbol{y}})\boldsymbol{v}_s - \left[\boldsymbol{h}^T(\vec{\boldsymbol{y}}) - \boldsymbol{g}^T(\vec{\boldsymbol{y}})\boldsymbol{Y}\right]\boldsymbol{p}. \tag{81}$$

Additional sources as discussed in the previous section are considered either as

$$p(\vec{\boldsymbol{y}}) \;=\; p^i(\vec{\boldsymbol{y}}) + \boldsymbol{g}^T(\vec{\boldsymbol{y}})\left[\boldsymbol{v}_s - \boldsymbol{v}_f^i\right] - \left[\boldsymbol{h}^T(\vec{\boldsymbol{y}}) - \boldsymbol{g}^T(\vec{\boldsymbol{y}})\boldsymbol{Y}\right]\boldsymbol{p} + \boldsymbol{h}^T(\vec{\boldsymbol{y}})\boldsymbol{p}^i \tag{82}$$

or as

$$p(\vec{\boldsymbol{y}}) \;=\; p^i(\vec{\boldsymbol{y}}) + \boldsymbol{g}^T(\vec{\boldsymbol{y}})\boldsymbol{v}_s - \left[\boldsymbol{h}^T(\vec{\boldsymbol{y}}) - \boldsymbol{g}^T(\vec{\boldsymbol{y}})\boldsymbol{Y}\right]\boldsymbol{p}. \tag{83}$$

Similar to Eqs. (66), (73), (82) and (83) are fully equivalent to each other.

4.2 Returning to the 1d Example

In the 1d example in one of the previous subsections, only the boundary data, i.e. p_0 and p_l were evaluated. It has been mentioned that this solution actually accounts for the solution of the boundary value problem which is a consequence of the fact that it has been the result of the only inversion in this process. However, field point evaluations are of substantial interest in many cases. In the specific case of the 1d duct problem, the field point sound pressure is evaluated using Eq. (81) with $\boldsymbol{p}$, $\boldsymbol{v}_s$

and $\boldsymbol{Y}$ according to Eqs. (51) and (54). The column matrices $\boldsymbol{g}$ and $\boldsymbol{h}$ are determined as

$$\begin{aligned} \boldsymbol{g}(y) &= -\frac{s}{2}\begin{bmatrix} \sin(ky) \\ \sin(k(l-y)) \end{bmatrix} \\ \boldsymbol{h}(y) &= -\frac{1}{2}\begin{bmatrix} \cos(ky) \\ \cos(k(l-y)) \end{bmatrix} \end{aligned} \tag{84}$$

Again, the 1d example shows that all boundary (element) matrices can even be set up in a 1d example. The 1d example is well suited to improve the understanding of the entire problem.

4.3 Rayleigh Integral and Other Methods to Approximate Radiated Sound Power

The Rayleigh integral is a simplification of the representation formula originally formulated for flat baffled radiators. (A baffle assumes symmetry conditions in the plane of the baffle.) In this case, the representation formula (75) simplifies to

$$p(\vec{\boldsymbol{y}}) = \int_{\Gamma} G(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}})\, sk v_f(\vec{\boldsymbol{x}}) d\Gamma(\vec{\boldsymbol{x}}) \tag{85}$$

because the normal derivative of the Green's function is always zero. Note that due to the symmetry, the Green's function here is twice the value of the one given in (12). For the case of the flat baffled radiator, the Rayleigh integral is exact. Furthermore, it is quite common to use the Rayleigh integral for other radiators as long as they are close to convex. However, in these cases, the Rayleigh integral is only used to approximate the radiated sound power P which is defined as

$$P = \frac{1}{2}\Re\left\{\int_{\Gamma_s} p v_f^* d\Gamma_s\right\} \tag{86}$$

where Γ_s is an arbitrary enveloping surface around the radiator. It is common to substitute Γ for Γ_s. The discretized version of the radiated sound power is written as

$$P = \frac{1}{2}\Re\left\{\boldsymbol{p}^T \boldsymbol{\Theta} \boldsymbol{v}_f^*\right\}. \tag{87}$$

It utilizes the boundary mass matrix $\boldsymbol{\Theta}$ again. Hence, evaluation of the radiated sound power requires solution of the acoustic boundary value problem. Since this is considered to be computationally too costly in an optimization process, a number of approximate methods have been developed. One of them is the well-known

approximation of the Rayleigh integral which is called the Lumped Parameter Model (LPM), cf. Koopmann and Fahnline (1997). It is compared with the accuracy of a full BEM solution and two other, even simpler, methods in Fritze et al. (2009). Therein, it is shown that in certain cases, the Rayleigh integral can be used to accurately approximate the radiated sound power.

5 Evaluation of Boundary Element System Matrices

5.1 General Remarks and Integral Free Term

In this section, the focus will be the evaluation of the boundary element system matrices $\boldsymbol{G}$ and $\boldsymbol{H}$ as given in Eqs. (25) and (26). The description will be limited to the collocation approach and assume sound hard boundary conditions, i.e. $Y = 0$. It will also be limited to three-dimensional problems where the boundary is the two-dimensional surface. At first, only quadrilaterals are addressed. Triangles will be discussed at the end of this section.

The integral free term in Eq. (26) requires to know the value of $c(\vec{\boldsymbol{y}})$ when $\vec{\boldsymbol{y}}$ is located on the boundary. Although possible to evaluate directly – it is related to the solid angle of this surface point – even for arbitrary corners and edges, it is efficient and more convenient to use an indirect approach. This implicit evaluation for an interior problem is based on the fact that when a uniform static pressure is applied, the particle velocity remains zero. Hence, for $k = 0$, it is $\boldsymbol{H}_{k=0}\boldsymbol{p} = \boldsymbol{0}$. A static and uniform pressure assumes that all entries of $\boldsymbol{p}$ are taking the same value. This implies that the sum of the coefficients of each row of $\boldsymbol{H}_{k=0}$ is zero so it can be concluded that

$$c(\vec{\boldsymbol{y}}) = -\int_{\Gamma} \frac{\partial G(k=0, \vec{\boldsymbol{x}}, \vec{\boldsymbol{y}})}{\partial n(\vec{\boldsymbol{x}})} d\Gamma(\vec{\boldsymbol{x}}). \tag{88}$$

This consideration is very closely related to the fact that pure Neumann problems of the Helmholtz equation will always provide one zero eigenvalue, similar to elastodynamic problems. Hence, the static stiffness matrix of a Neumann problem must be singular. Note that this is only valid for closed domains, i.e. interior problems. For exterior problems, the complementary value of 1 is required. Therefore, c is evaluated as

$$c(\vec{\boldsymbol{y}}) = 1 - \int_{\Gamma} \frac{\partial G(k=0, \vec{\boldsymbol{x}}, \vec{\boldsymbol{y}})}{\partial n(\vec{\boldsymbol{x}})} d\Gamma(\vec{\boldsymbol{x}}) \tag{89}$$

if an exterior problem is considered.

It is the experience of the author that evaluation of these integral free terms for the geometric nodes of the mesh is a very useful model check. A complex boundary element mesh is not always easy to survey and sometimes there are elements with a wrong definition of the normal vector. Sometimes, there are overlaps or gaps in the surface mesh. In all these cases, the user will find irregularities in the values of c and

is easily alerted. Hence, evaluation of these terms is even recommended if – such as for discontinuous elements – all collocation points are located within the elements and thus $c = 0.5$.

5.2 Numerical Determination of Regular Integrals

The collocation method as described here requires an integration over the entire surface for each collocation point. Therefore, the collocation point $\vec{z}_l$ in (25) and (26) defines the row of either matrix, $\boldsymbol{G}$ and $\boldsymbol{H}$. The parameter which is used for these integrations is the Euclidean distance $r(\vec{x}, \vec{z}_l)$ between the collocation point and the surface. This is tried to be depicted in Fig. 2 which is sketching a 2d domain for simplicity only. Since integration is usually performed on the element level. This is the reason why, in what follows, only one collocation point (node) and one element are considered. Note that when using continuous elements, one basis function will contribute to more than one element. When writing a BEM code, it is a straightforward approach to use an outer loop over all nodes and inner loop over all elements.

It is usually impossible or not practical to integrate over the element analytically. For numeric integration, it is convenient to introduce a coordinate transformation such that further steps are carried out on a reference element for which very efficient integration techniques have been developed. Herein, the reference element is defined by coordinates η_1 and η_2 in the interval $[-1, 1]$. At first, it is necessary to formulate the position vector in terms of these coordinates, see Fig. 3.

It is common and, of course, useful to approximate the element geometry by using Lagrangian polynomials N_k. Thus the position vector becomes a function as $\vec{\boldsymbol{x}} = \vec{\boldsymbol{x}}(N_k(\eta_1, \eta_2), \vec{\boldsymbol{P}}_k)$. Lagrangian polynomials possess the property that they are one at only one distinct node and zero at all other nodes of the element as $N_k \vec{\boldsymbol{P}}_k = \vec{\boldsymbol{P}}_k \delta_{kl}$. Hence, we can write

$$\vec{\boldsymbol{x}} = \sum N_k(\eta_1, \eta_2) \vec{\boldsymbol{P}}_k. \tag{90}$$

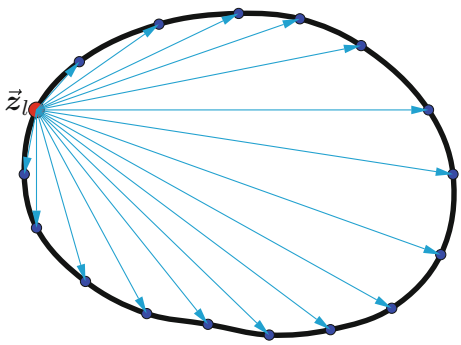


Fig. 2 Vivid description of the integration over the entire boundary. The integral kernel depends on the Euclidean distance between the collocation point and the remaining surface

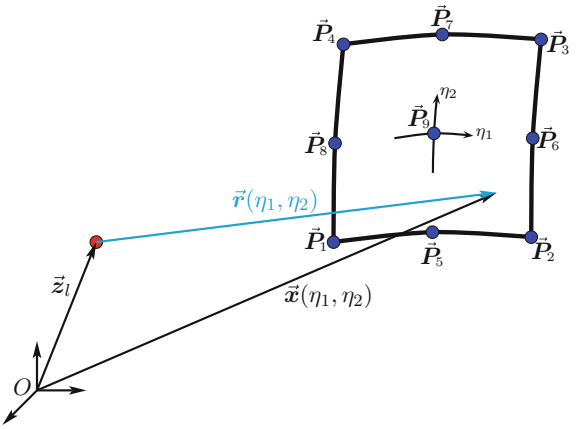


Fig. 3 Configuration of collocation point and element for definition of the position vector $\vec{x}$ and the distance vector $\vec{r}$ with numbering of geometric nodes

In literature, these functions N_k are often called shape functions because they approximate the geometry, i.e. the shape. However, many authors mix this with the interpolation functions which are used for the physical quantity. This can create substantial confusion and should be avoided. In particular in boundary element methods, it is useful to clearly distinguish between geometry approximation (shape approximation) and interpolation of the quantity which the boundary value problem is to be solved for, i.e. in our case the sound pressure.

In case of a linear geometry approximation where only the nodes $\vec{P}_1 - \vec{P}_4$ are used, the position vector is written as

$$\begin{aligned}\vec{x} = \quad & \frac{1}{4}(1-\eta_1)(1-\eta_2)\,\vec{P}_1 + \frac{1}{4}(1+\eta_1)(1-\eta_2)\,\vec{P}_2 + \\ & + \frac{1}{4}(1+\eta_1)(1+\eta_2)\,\vec{P}_3 + \frac{1}{4}(1-\eta_1)(1+\eta_2)\,\vec{P}_4\end{aligned} \tag{91}$$

whereas the quadratic shape approximation is utilizing nine nodes and is written as

$$\begin{aligned}\vec{x} = & \frac{1}{4}\eta_1(1-\eta_1)\,\eta_2(1-\eta_2)\,\vec{P}_1 + \frac{1}{4}\eta_1(1+\eta_1)\,\eta_2(1-\eta_2)\,\vec{P}_2 + \\ & + \frac{1}{4}\eta_1(1+\eta_1)\,\eta_2(1+\eta_2)\,\vec{P}_3 + \frac{1}{4}\eta_1(1-\eta_1)\,\eta_2(1+\eta_2)\,\vec{P}_4 + \\ & - \frac{1}{2}\left(1-\eta_1^2\right)\eta_2(1-\eta_2)\,\vec{P}_5 + \frac{1}{2}\eta_1(1+\eta_1)\left(1-\eta_2^2\right)\vec{P}_6 + \\ & + \frac{1}{2}\left(1-\eta_1^2\right)\eta_2(1+\eta_2)\,\vec{P}_7 - \frac{1}{2}\eta_1(1-\eta_1)\left(1-\eta_2^2\right)\vec{P}_8 + \\ & + \left(1-\eta_1^2\right)\left(1-\eta_2^2\right)\vec{P}_9.\end{aligned} \tag{92}$$

It is easy to determine the Euclidean distance and the distance vector as the difference between $\vec{x}$ and $\vec{z}_l$.

Although not common, it is convenient to use tensor algebra calculus to determine the Jacobian determinant for curved coordinates. For this, the covariant basis consisting of the basis vectors is formulated. These basis vectors $\vec{g}_1$ and $\vec{g}_2$ can be understood as tangential vectors to the coordinate curves which are defined by constant values of η_2 and η_1, respectively. They are determined by

$$\vec{g}_k = \frac{\partial \vec{x}}{\partial \eta_k}. \tag{93}$$

The covariant basis vectors form the metric tensor $\vec{\vec{g}}$ by setting up their scalar products as

$$\vec{\vec{g}} = \begin{bmatrix} \vec{g}_1 \cdot \vec{g}_1 & \vec{g}_1 \cdot \vec{g}_2 \\ \vec{g}_2 \cdot \vec{g}_1 & \vec{g}_2 \cdot \vec{g}_2 \end{bmatrix} = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix}. \tag{94}$$

The Jacobi determinant is then determined as the square root of the Gram determinant

$$J = \sqrt{\det\left(\vec{\vec{g}}\right)} = \sqrt{g_{11}g_{22} - g_{12}g_{21}}. \tag{95}$$

This calculus is convenient to determine the unit normal vector on the element. Note that on a curved element the normal vector varies with location. The normal vector is normal with respect to the tangential basis as

$$\vec{n} = \frac{\vec{g}_1 \times \vec{g}_2}{J}. \tag{96}$$

For somebody writing a BEM code, it is important to develop a strategy how to guarantee that the normal vector is always pointing into Ω_c. This must be controlled somehow and is possible by controlling the order of nodes such that they are always counted in the mathematically positive direction, i.e. anticlockwise.

Coming back to the integrals, it can be stated that integrals of two forms are to be computed, i.e.

$$\begin{aligned} I_G &= \int_{\Gamma_e} G(\vec{x}, \vec{y})\varphi(\vec{x}) d\Gamma_e(\vec{x}) \\ &= \int_{-1}^{1}\int_{-1}^{1} G(\vec{x}(\eta_1, \eta_2), \vec{y})\varphi(\eta_1, \eta_2) J(\eta_1, \eta_2) d\eta_2 d\eta_1 \end{aligned} \tag{97}$$

and

$$
\begin{aligned}
I_H &= \int_{\Gamma_e} \frac{\partial G(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}})}{\partial n(\vec{\boldsymbol{x}})} d\Gamma(\vec{\boldsymbol{x}}) \\
&= \int_{-1}^{1}\int_{-1}^{1} \frac{\partial G(\vec{\boldsymbol{x}}(\eta_1,\eta_2), \vec{\boldsymbol{y}})}{\partial n(\vec{\boldsymbol{x}}(\eta_1,\eta_2))} \varphi(\eta_1,\eta_2) J(\eta_1,\eta_2) d\eta_2 d\eta_1,
\end{aligned}
\tag{98}
$$

where

$$
\begin{aligned}
G(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}}) &= \frac{1}{4\pi} \frac{\mathrm{e}^{\mathrm{i}kr(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}})}}{r(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}})} \\
\frac{\partial G(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}})}{\partial n(\vec{\boldsymbol{x}})} &= -\frac{1}{4\pi} \frac{1 - \mathrm{i}kr(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}})}{r^2(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}})} \mathrm{e}^{\mathrm{i}kr(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}})} \frac{\partial r(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}})}{\partial n(\vec{\boldsymbol{x}})}
\end{aligned}
\tag{99}
$$

and

$$
\frac{\partial r}{\partial n} = \vec{\boldsymbol{\nabla}} r \cdot \vec{\boldsymbol{n}} = \frac{\partial r}{\partial x_i} n_i, \tag{100}
$$

with summation convention applied at the right end. Analysing these integrals, two categories of kernel functions are identified. As such, these integrals are written as

$$
\begin{aligned}
I_G &= \int_{-1}^{1}\int_{-1}^{1} \bar{f}(\eta_1,\eta_2) \frac{\mathrm{e}^{\mathrm{i}kr}}{r} J(\eta_1,\eta_2) d\eta_2 d\eta_1 \\
I_H &= \int_{-1}^{1}\int_{-1}^{1} \bar{f}(\eta_1,\eta_2) \frac{\mathrm{e}^{\mathrm{i}kr}}{r^2} \frac{\partial r}{\partial n} J(\eta_1,\eta_2) d\eta_2 d\eta_1,
\end{aligned}
\tag{101}
$$

with $r = r(\eta_1, \eta_2)$ and $\bar{f}(\eta_1, \eta_2)$ being a smooth function behaving like a polynomial of a certain order. The function $\mathrm{e}^{\mathrm{i}kr}$ is an oscillating function. However, in comparison with $\bar{f}$, it simply adds some oscillation. If the size of the element is much smaller than the wavelength, this term behaves similar to a low order polynomial, whereas the polynomial in $\bar{f}$ should be able to approximate this term. Similarly, J is (usually) a rather smooth function. Hence, it can be concluded that the function $f = \bar{f}\mathrm{e}^{\mathrm{i}kr} J$ behaves like a polynomial of doubled (or maybe tripled) order in comparison with $\bar{f}$. Keeping in mind that f is a function of, at least, double order of the interpolation polynomials φ, the integrals can be rewritten as

$$
\begin{aligned}
I_G &= \int_{-1}^{1}\int_{-1}^{1} f(\eta_1,\eta_2) \frac{1}{r} d\eta_2 d\eta_1 \\
I_H &= \int_{-1}^{1}\int_{-1}^{1} f(\eta_1,\eta_2) \frac{1}{r^2} \frac{\partial r}{\partial n} d\eta_2 d\eta_1.
\end{aligned}
\tag{102}
$$

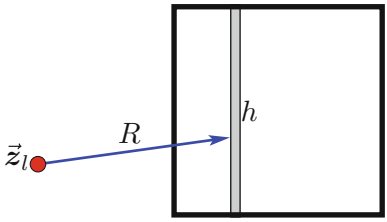


Fig. 4 Configuration of distance R and element size h for evaluation of the decision criterion D in Eq. (104)

Since these integrals are evaluated numerically, it is time to take a look at numerical integration techniques. Because it provides exponential convergence, Gauss–Legendre quadrature is a well suited technique for numerical integration in this case. This means that an arbitrary function g is approximately integrated as

$$\int_{-1}^{1}\int_{-1}^{1} g(\eta_1,\eta_2)d\eta_2 d\eta_1 = \sum_{i=1}^{n_i}\sum_{k=1}^{n_k} g(\eta_{1_i},\eta_{2_k})w_i w_k, \tag{103}$$

where η_{m_j} are the zeros of the Legendre polynomials, i.e. Gauss points, and w_j are the associated weights.

It is of crucial importance to the efficiency of the entire simulation that it is a priori clear how many integration points are used in the process. Although suitable suggestions are found in early literature on BEM, cf. Brebbia et al. (1984), this topic does not seem to be well understood among the users of BEM in acoustics. The author of this paper has realized that if people even talk about their integration rules, it seems to be common that – mostly – only singular and regular integrals are distinguished. Of course, this is not enough. The diminishing gradient of the function $1/r$ and theoretically (but not in practice) $1/r^2$ require an adaptive integration scheme. This can be based on a distance function. Note that the function $1/r$ is smoothing as r becomes larger. A very suitable function to decide the order of integration was proposed in the computer codes by Brebbia et al. (1984). This decision criterion considers the ratio between distance of the collocation point from the element and size of the element as

$$D = \frac{2R}{h}. \tag{104}$$

A vivid description of this criterion is shown in Fig. 4.

It is not easy for the author to give reliable and close to the edge suggestions on the order of integration. It is, however, the experience of the author that in the far field, i.e. $D > 20$, it is sufficient to use between 2 and 4 integration points per direction. This assumes that interpolation polynomials of order ≤ 2 are used. It is rather clear that constant elements require only an integration of 2 (per direction) while quadratic elements should be integrated one order higher, maybe even with fourth order to be safe. The author observed that, in large scale models of more than 20000 degrees of freedom, the number of integrals with $D > 20$ is, at least, two orders of magnitude larger than all other integrals. Note that increasing the order of integration from 2 to

4 per direction increases the computational costs by factor 4, from 3 to 4 even by a factor of almost 2. If D is decreasing, the order of integration should be increased. However, the author is not aware of any sufficient investigation of how this increase should be quantified and, thus is trying to remain on the safe side by substantially increasing the order of integration as D becomes smaller.

When D is getting small (but not zero), we speak of quasi-singular or nearly singular integrals. Numerical evaluation of these integrals suffers from the large gradient which is due to the $1/r$ or $1/r^2$ function. For these types of integrals, a nice third order polynomial transformation technique has been proposed to efficiently evaluate these integrals, cf. Telles (1987). In what follows, the algorithm will be presented as a recipe. The third order polynomial transformation shifts the location of the Gauss points to adjust them better to the quasi-singularity. It is applied to both directions on the quadrilateral element independently. It is assumed that $\tilde{\eta}$ is the original Gauss point, i.e. a zero of a Legendre polynomial, and η_s is the position on the element, which is the closest to the singularity, i.e. usually the collocation point in the vicinity of the element under consideration. Then, the new location of the Gauss point η is determined by

$$\eta = a\tilde{\eta}^3 + b\tilde{\eta}^2 + c\tilde{\eta} + d \tag{105}$$

with the additional Jacobian J_a as

$$J_a = \frac{d\eta}{d\tilde{\eta}} = 3a\tilde{\eta}^2 + 2b\tilde{\eta} + c. \tag{106}$$

The parameters of this transformation are given as

$$a = \frac{1-\bar{R}}{1-3\gamma^2}, \qquad b = -\frac{3\gamma(1-\bar{R})}{1-3\gamma^2}, \qquad c = \frac{\bar{R}+3\gamma^2}{1-3\gamma^2}, \qquad d = -b \tag{107}$$

where

$$\gamma = \sqrt[3]{-q+\sqrt{q^2+p^3}} - \sqrt[3]{q+\sqrt{q^2+p^3}} + \frac{\eta_s}{1+2\bar{R}} \tag{108}$$

with

$$\begin{aligned} q &= \frac{\eta_s}{(1+2\bar{R})^3}\left[1-4\bar{R}^2-\eta_s^2\right] \\ p &= \frac{1}{3(1+2\bar{R})^2}\left[4\bar{R}\left(1-\bar{R}\right)+3\left(1-\eta_s^2\right)\right]. \end{aligned} \tag{109}$$

Finally, the transformation depends on two parameters of which η_s is determined as mentioned above. Thus, the only parameter remaining is the value of $\bar{R}$. In Telles

Table 1 Number of integrals evaluated for certain values of D in four different boundary element models

Model	Cube	Radiatterer		
Interpolation	P_{1d}	P_0	P_{1d}	P_{2c}
Elements	7776	18216	4554	4554
Nodes	31104	18216	18216	18214
Interval				
$D \geq 20$	459480192	636689116	129836376	129938746
$20 > D \geq 12$	15132144	19150228	24711080	24716880
$12 > D \geq 7$	5625984	5854804	7919636	7960938
$7 > D \geq 3.618$	2231088	1174020	2633336	2491712
$3.618 > D \geq 1.2$	1011216	594984	664764	593160
$1.2 > D \geq 0.4$	186576	145728	109704	109704
$0.4 > D \geq 0.1$	0	0	0	0
$0.1 > D \geq 0.01$	0	0	0	0

(1987), this parameter is chosen after an optimization procedure. It had been proposed to chose it as

$$\begin{aligned} \bar{R} &= 0.85 + 0.24 \ln D && \text{for} \quad 0.05 \leq D \leq 1.3 \\ \bar{R} &= 0.893 + 0.0832 \ln D && \text{for} \quad 1.3 \leq D \leq 3.618 \\ \bar{R} &= 1 && \text{for} \quad D \geq 3.618. \end{aligned} \tag{110}$$

Note that $\bar{R} = 1$ means that there is no transformation.

Although being clear that for interpolation functions of order zero, i.e. constant elements, to order two, i.e. quadratic elements a minimum order of 2–3 (per direction) should be used for $D > 20$, it is difficult to recommend a certain order for smaller values of D. The author is increasing the order in certain steps with decreasing D. However, the actual order of integration will be driven by the accuracy of the analysis. Hence, in some cases it may be fine to use lower order, in some cases higher order integration is recommended. In the example of the boundary element model of a cube and another example of the Radiatterer, cf. Hornikx et al. (2015) and Marburg (2016a), the author has counted the number of integrals in certain intervals of D, see Table 1. There are meshes of constant elements P_0, linear discontinuous elements P_{1d} and continuous quadratic elements P_{2c}. In all cases, the meshes are very regular in that all elements are actually shaped as squares of the same size. It can be seen from these data that the majority of integrals of a large scale model occurs for large numbers of D. Very small values of D occur in irregular meshes and in cases where thin bodies or gaps are modelled. Special care should be taken then.

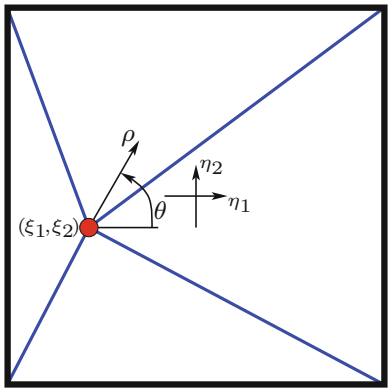


Fig. 5 Configuration with singularity at (ξ_1, ξ_2) on the element

5.3 Singular Integrals

It is one of the properties of the boundary element method that singular integrals occur and need to be evaluated. A singular integral occurs if the collocation point is part of the element under consideration. In particular, the early literature on boundary element methods is full of techniques to deal with these integrals. Herein, only one strategy is presented. It is based on a coordinate transformation using polar coordinates. Although described in many papers, the author is presenting this technique mainly by using the same description as do Rego Silva (1993).

Figure 5 shows the configuration for an element with the collocation point at $\eta_1 = \xi_1$ and $\eta_2 = \xi_2$. Again, the integrals to be evaluated are the same as given in Eq. (102)

$$
\begin{aligned}
I_G &= \int_{-1}^{1}\int_{-1}^{1} f(\eta_1, \eta_2)\frac{1}{r}d\eta_2 d\eta_1 \\
I_H &= \int_{-1}^{1}\int_{-1}^{1} f(\eta_1, \eta_2)\frac{1}{r^2}\frac{\partial r}{\partial n}d\eta_2 d\eta_1.
\end{aligned}
\tag{111}
$$

It is obvious that the kernel function of I_G goes to infinity of $r \to 0$. Such an integral is called weakly singular. Although the literature is full of remarks that I_H would be an integral in the sense of Cauchy's principal value, it must be mentioned that it is not. In fact, it is a regular integral, cf. Kirkup (1998).

To be able to integrate over the weak singularity in I_G, it is useful to introduce the coordinate transformation as

$$
\begin{aligned}
\eta_1 &= \xi_1 + \rho\cos\theta \\
\eta_2 &= \xi_2 + \rho\sin\theta \\
d\eta_1 d\eta_2 &= \rho d\rho d\theta
\end{aligned}
\tag{112}
$$

where the polar coordinates ρ and θ are introduced and the origin of this system is put at the singularity. This modifies the integrals I_G and I_H into

$$I_G = \sum_{t=1}^{4} \int_{\theta_1}^{\theta_2} \int_0^{\hat{\rho}(\theta)} f(\rho, \theta) \frac{1}{r(\rho)} \rho d\rho d\theta$$
$$I_H = \sum_{t=1}^{4} \int_{\theta_1}^{\theta_2} \int_0^{\hat{\rho}(\theta)} f(\rho, \theta) \frac{1}{r^2(\rho)} \frac{\partial r(\rho)}{\partial n} \rho d\rho d\theta \tag{113}$$

which means that they are now split into four integrals over triangles. The upper limit $\hat{\rho}(\theta)$ can be determined by dividing the horizontal or vertical distance from the new origin to the element edge by $\cos\theta$ or $\sin\theta$, respectively. The Euclidian distance r is the same as ρ for a flat element. Then, the singularity cancels out for I_G and is not there at all for I_H, since the normal derivative is zero on a flat element. Explanation of the regularity of the integrals in Eq. (113) is more advanced for curved elements and beyond the scope of this chapter. However, even then these integrals are easily evaluated by using standard quadrature rules.

5.4 Triangular Elements

So far, the coordinate transformation and the integration techniques were explained for quadrilateral elements. It is easy to extend these techniques to triangles by introducing another coordinate transformation. This is shown in Fig. 6. Often, triangular coordinates γ_1, γ_2 are defined in the interval [0, 1]. The integrals can be transformed as follows

$$\int_0^1 \int_0^{1-\gamma_2} f(\gamma_1, \gamma_2) d\gamma_1 d\gamma_2 = \int_{-1}^1 \int_{-1}^1 f(\eta_1, \eta_2)\,(\eta_1 + 1)\, d\eta_2 d\eta_1 \tag{114}$$

and the triangular coordinates γ_j are related to η_j by

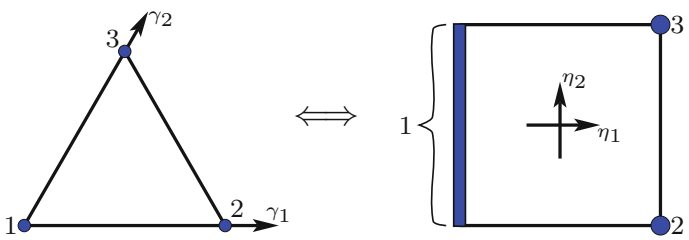


Fig. 6 Transformation of triangular coordinates to coordinates of a quadrilateral

$$\gamma_1 = \frac{1}{4}(1+\eta_1)(1-\eta_2) \qquad \gamma_2 = \frac{1}{4}(1+\eta_1)(1+\eta_2)$$
$$\eta_1 = 2(\gamma_1+\gamma_2) - 1 \qquad \eta_2 = \frac{\gamma_2-\gamma_1}{\gamma_2+\gamma_1}. \tag{115}$$

By this transformation, the triangular element is mapped into a quadrilateral with the corner $\gamma_1 = \gamma_2 = 0$ being transformed into the side $\eta_1 = -1$. Experiences with this transformation are quite positive as the integration is performed very reliably and accurately.

6 Choice of Boundary Elements

6.1 General Remarks

The choice and performance of boundary elements has been described and discussed in the book chapter by the author, cf. Marburg (2008), see also the papers Marburg (2002b) and Marburg and Schneider (2003b). These results will be briefly summarized in this section.

Since the collocation boundary element method seems to perform the best with Lagrangian elements, i.e. elements for which Lagrangian polynomials are used for interpolation, this type of elements will be discussed only. Among the Lagrangian elements, it is possible to distinguish continuous and discontinuous elements. Different from conventional finite element methods, boundary element methods do not require continuity of the physical quantity over the element's boundary. This allows not only to build models which have so-called hanging nodes, i.e. nodes at the edge of two elements but only used by one of them, see for example Fig. 7. It even allows elements with nodes inside the element and not at the element's edge as shown in Fig. 8. This will be discussed in more detail in this section.

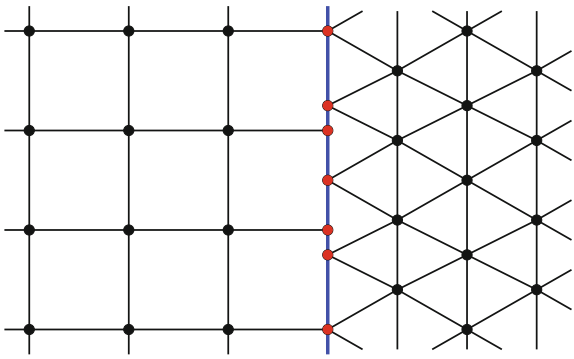


Fig. 7 Possible mesh for BEM: discontinuous with respect to physical quantities, geometry must be continuous

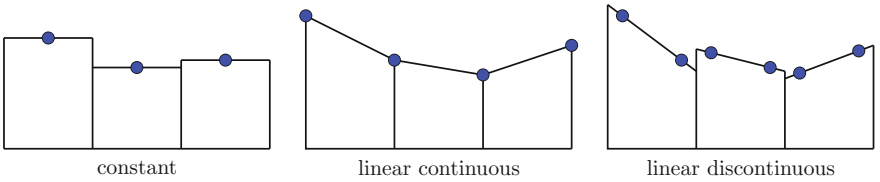


Fig. 8 Vivid description of discontinuous and continuous element approximation of physical quantities

Another question in this context regards mesh size. It seems to be widely accepted that the numerical error of boundary element methods is controlled by the number of elements per wavelength. This was investigated in the author's studies mentioned above and will also be discussed in what follows.

6.2 Continuous Boundary Elements

Continuous elements account for the most commonly used types of boundary elements. Most likely, this is due to the experiences which users have gained in the context of finite elements and, also, due to the misunderstanding that a continuous physical quantity, e.g. the sound pressure, should be approximated by continuous functions. The most popular elements are constructed by (bi)linear and (bi)quadratic Lagrangian interpolation functions.

Interpolation functions of continuous quadrilateral surface elements are easily constructed by multiplying two one-dimensional polynomials ψ_1 and ψ_2, cf. do Rego Silva (1993). Introducing the notation of upper indices l and q for linear and quadratic polynomials, respectively, these linear polynomials are formulated as

$$\begin{aligned} \psi_1^l(\eta_k) &= \frac{1}{2}(1-\eta_k) \qquad \text{and} \\ \psi_2^l(\eta_k) &= \frac{1}{2}(1+\eta_k)\,, \end{aligned} \tag{116}$$

whereas quadratic polynomials are given by

$$\begin{aligned} \psi_1^q(\eta_k) &= -\frac{1}{2}\eta_k(1-\eta_k) \quad , \\ \psi_2^q(\eta_k) &= \frac{1}{2}\eta_k(1+\eta_k) \quad \text{and} \\ \psi_3^q(\eta_k) &= (1-\eta_k^2) \quad . \end{aligned} \tag{117}$$

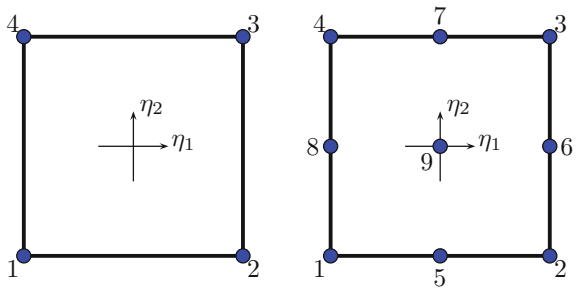


Fig. 9 Configuration of a linear (*left*) boundary element and a quadratic (*right*) boundary element

The actual interpolation functions φ_l on the quadrilateral element are evaluated by multiplying the two one-dimensional polynomials $\psi_i(\eta_1)$ and $\psi_k(\eta_2)$. Relating this to the elements in Fig. 9 yields

$$\begin{aligned} \varphi_1^l &= \psi_1^l(\eta_1)\,\psi_1^l(\eta_2)\ ,\ \varphi_2^l = \psi_2^l(\eta_1)\,\psi_1^l(\eta_2)\ , \\ \varphi_3^l &= \psi_2^l(\eta_1)\,\psi_2^l(\eta_2)\ ,\ \varphi_4^l = \psi_1^l(\eta_1)\,\psi_2^l(\eta_2) \end{aligned} \tag{118}$$

and

$$\begin{aligned} &\varphi_1^q = \psi_1^q(\eta_1)\,\psi_1^q(\eta_2)\,, \quad \varphi_2^q = \psi_2^q(\eta_1)\,\psi_1^q(\eta_2)\,, \quad \varphi_3^q = \psi_2^q(\eta_1)\,\psi_2^q(\eta_2)\,, \\ &\varphi_4^q = \psi_1^q(\eta_1)\,\psi_2^q(\eta_2)\,, \quad \varphi_5^q = \psi_3^q(\eta_1)\,\psi_1^q(\eta_2)\,, \quad \varphi_6^q = \psi_2^q(\eta_1)\,\psi_3^q(\eta_2)\,, \\ &\varphi_7^q = \psi_3^q(\eta_1)\,\psi_2^q(\eta_2)\,, \quad \varphi_8^q = \psi_1^q(\eta_1)\,\psi_3^q(\eta_2)\,, \quad \varphi_9^q = \psi_3^q(\eta_1)\,\psi_3^q(\eta_2)\,. \end{aligned} \tag{119}$$

Finally, the interpolation functions φ_j are more or less the same as those applied for the geometry approximation in Eqs. (91) for linear and (92) for quadratic approximation.

In what follows, linear continuous elements will be referred to as P_{1c} whereas quadratic continuous elements will be called P_{2c}. (This has been used in the previous section already.)

6.3 Discontinuous Boundary Elements

An alternative to the commonly used continuous elements consists in the use of discontinuous elements. For discontinuous elements, it is impossible that the approximation of the geometry coincides with the interpolation of the physical quantity. While the geometry is approximated with at least a certain number of points at the element edge, the interpolation and collocation points are located inside the element as shown in Figs. 8 and 10.

Interpolation functions of discontinuous quadrilateral elements are constructed in a similar way as the interpolation functions of continuous elements. The simplest

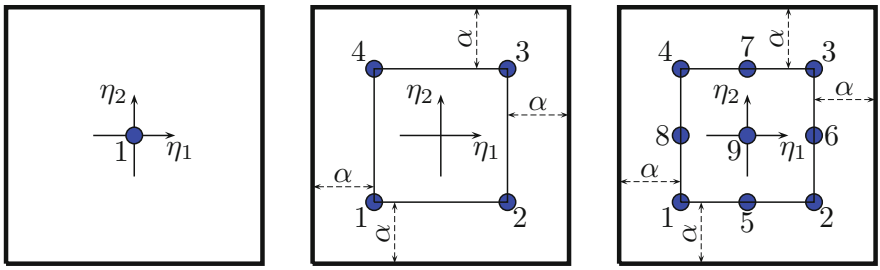


Fig. 10 Configuration of discontinuous constant (*left*), linear (*center*) and quadratic (*right*) boundary elements

discontinuous elements use constant interpolation. Hence, we write the constant interpolation function as

$$\psi_1^c(\eta_k) = 1 \quad . \tag{120}$$

For linear and quadratic discontinuous elements, we assume that the distance between the element edge and the closest nodal point on the standard element is given by the value of α with $0 < \alpha < 1$. Introducing the constant $\zeta = 1 - \alpha$, we write

$$\begin{aligned} \psi_1^l(\eta_k) &= \frac{1}{2\zeta}\,(\zeta - \eta_k) \qquad \text{and} \\ \psi_2^l(\eta_k) &= \frac{1}{2\zeta}\,(\zeta + \eta_k)\ , \end{aligned} \tag{121}$$

whereas quadratic polynomials are given by

$$\begin{aligned} \psi_1^q(\eta_k) &= \frac{1}{2\zeta^2}\,\eta_k\,(\eta_k - \zeta) \qquad , \\ \psi_2^q(\eta_k) &= \frac{1}{2\zeta^2}\,\eta_k\,(\eta_k + \zeta) \qquad \text{and} \\ \psi_3^q(\eta_k) &= \frac{1}{\zeta^2}\,(\zeta - \eta_k)\,(\zeta + \eta_k) \quad . \end{aligned} \tag{122}$$

The actual interpolation functions φ are again evaluated by multiplying the two one-dimensional functions. In case of the constant elements, this is simple as

$$\varphi_1^c = \psi_1^c(\eta_1)\psi_2^c(\eta_2) = 1 \quad . \tag{123}$$

For linear and for quadratic elements, Eqs. (118) and (119) are applied, respectively.

The open parameter of discontinuous elements as described here consists in the value of α. The straightforward approach would chose α such that, in a regular mesh, all collocation points are located equidistant from each other. Then we have $\alpha = 0.5$

for linear and $\alpha = 0.3333$ for quadratic elements. Alternatively, the position may be chosen at the zeros of orthogonal polynomials, in particular the zeros of the Legendre polynomials, cf. Marburg (2008) and Marburg and Schneider (2003b). Zeros of the Legendre polynomials demand $\alpha = 0.4226$ and $\alpha = 0.2254$ for linear and quadratic quadrilaterals, respectively.

Similar to the continuous elements, linear discontinuous elements will be referred to as P_{1e} and P_{1L} for the collocation points with equidistant spacing or at the zeros of the Legendre polynomials, respectively, whereas quadratic discontinuous elements will be called P_{2e} and P_{2L} for the same reasons. Constant elements which cannot be continuous at all receive the notation of P_0.

For consideration of triangles, we refer to the literature, again cf. Marburg (2008) and Marburg and Schneider (2003b).

6.4 Error Measures

It is necessary for comparison of different elements (or methods) that error measures are defined. These error measures are functions depending on parameters such as position and frequency and are usually based on a comparison of the current solution with a reference model. In this chapter, the examples under consideration will be a duct and a sedan cabin. Both examples have already been discussed in detail in the author's work, cf. Marburg (2002b), Marburg and Schneider (2003b), Marburg (2008). It is a common approach for testing numerical models that the reference solution is either chosen as the analytical solution of the problem or as a so-called overkill solution, i.e. a numerical model which is able to produce much more accurate results than the model currently tested. In more practical cases, experimental results can be used as reference. However, the researcher should always be aware that experimental results are always containing a certain measurement error which needs to be estimated as well.

An error function e^{Γ} is defined to measure the surface error as

$$e^{\Gamma}(\vec{x}) = \bar{p}(\vec{x}) - p(\vec{x}) \qquad \vec{x} \in \Gamma \tag{124}$$

where $\bar{p}(\vec{x})$ represents the approximate solution yielded by using the boundary element formulation and $p(\vec{x})$ represents the reference solution. In the case of the duct, the analytic solution of the one-dimensional duct problem accounts for the reference solution. In case of the sedan cabin, the reference solution is an overkill solution which is obtained by the finest discretization using quadratic elements P_{2L}. An error e^{Ω} is defined analogously in the interior domain, i.e. for $\vec{x} \in \Omega$, where the solution obtained at a number of discrete field points is compared with the reference solution.

The discrete error function is evaluated at discrete points, i.e. all collocation points for the surface error and a certain number of interior points for the error in the cavity. Then, the discrete surface error is determined as

$$||e^{\Gamma}||_m \;=\; \left(\frac{1}{N_n}\sum_{i=1}^{N_n}\|e(\vec{\boldsymbol{x}}_i)\|^m\right)^{\frac{1}{m}} \tag{125}$$

where N_n represents the number of nodes and m the specific norm, the Euclidean norm (rms) for $m = 2$ and the maximum norm for $m \to \infty$. The examples of the current section present only errors measured in the Euclidean norm.

In what follows, we usually use relative errors e_m^{Γ} for the sound pressure error

$$e_m^{\Gamma} \;=\; \frac{||e^{\Gamma}||_m}{||p^{\Gamma}||_m} \tag{126}$$

where $||p^{\Gamma}||_m$ accounts for the discrete norm of the exact sound pressure. Analogously, e_m^{Ω} accounts for the sound pressure error at a certain number of field points.

6.5 Computational Examples

Traveling waves in a long duct: The long duct, i.e. an air-filled duct of length $l = 3.4\,\mathrm{m}$ and a square cross section of $0.2 \times 0.2\,\mathrm{m}^2$ is a well suited test case since the three-dimensional numerical solution can be compared to the one-dimensional analytical solution, at least up to a frequency where modes perpendicular to the length of the duct occur. Assuming a speed of sound of 340 m/s and an ambient density of $1.3\,\mathrm{kg/m}^3$, these perpendicular modes are observed at a frequency of 1700 Hz. A solution with traveling plane waves is found if a particle velocity is applied to one end, i.e. an inhomogeneous Neumann boundary condition at $x = 0$, and a fully absorbing boundary condition is applied to the other end at $x = l$. Full absorption is achieved by an impedance of $Z = \rho c$. This is equivalent to a normalized boundary admittance of $\tilde{Y} = 1$. All other walls are considered acoustically rigid and at rest, i.e. a homogeneous Neumann condition is used. This configuration leads to a solution of constant sound pressure magnitude everywhere in the duct and at all frequencies. Only the phase angle varies. The solution of the one-dimensional problem is given as

$$p(x) \;=\; -v_s(0)\,\rho c\, e^{\mathrm{i}kx} \quad . \tag{127}$$

Independence of the sound pressure magnitude from position and frequency makes this example an ideally suited one to compare different element types. In what follows, the abbreviations P_0, P_{1c}, P_{2e} etc. will be used as introduced in the previous subsections.

Error dependence in terms of the element size is presented for two different frequencies in Fig. 11. It is easy to realize that different functions of error occur for different frequencies. For 500 Hz, lines for P_0, P_{1e}, P_{1c} and P_{1L} are almost parallel but on different levels. The remaining three functions are almost parallel too but much steeper. For 1500 Hz, the lines for P_0, P_{1e} and P_{1c} are nearly parallel. The

error for elements P_{1L}, however, is now parallel to the lines of quadratic elements which indicates higher convergence rate for these higher frequencies. Note, that the functions of error for P_{1L} and P_{2e} coincide!

So, it is realized that slopes of error for P_{1L} are about the same as for other linear or for constant elements at the low frequency of 500 Hz but much greater for the higher frequency of 1500 Hz. A similar behaviour is assumed for quadratic elements P_{2L}. The frequency is not large enough to confirm. (Solutions at higher frequencies are perturbed due to the ill-conditioning of perpendicular modes.)

Since the error functions in terms of mesh size show the same slopes at least for the P_0, P_{1e}, P_{1c}, P_{2c} and P_{2e} elements, it is assumed that an occurrence of a pollution effect which is well-known from finite elements, cf. Ihlenburg (1998) is very unlikely. Later in this subsection, it will be discussed why the slopes of error functions of P_{1L} and P_{2L} elements show or may show a dependence on frequency.

The abscissa is indicating in addition to the element size how many boundary elements per wave are used. At 500 Hz, this number lies between 3.4 and 27.2 whereas at 1500 Hz only a third of these is counted. It clearly shows that errors below 1% are only realistically reached when using discontinuous linear (or quadratic) elements or when using continuous quadratic elements. Constant and linear elements are converging slowly where the constant elements are still providing lower errors than the linear elements. As it is known from finite elements, it is necessary to use higher order elements to achieve a very low error, cf. Thompson and Pinsky (1994).

Although not shown here, it is mentioned that the situation is very similar if the error is measured in the maximum norm, cf. Marburg and Schneider (2003b). Remarkable differences are only observed for discontinuous elements P_{1L} and P_{2L}. This will be discussed later.

While the Fig. 11 showed the error in terms of the element size, it is usually of more practical relevance to know the error in terms of the degree of freedom. For large scale models, it is quite often the case that the degree of freedom controls the computational complexity, i.e. memory requirements and computation time. For a certain number N_e of quadrilateral boundary elements, we find that the number of

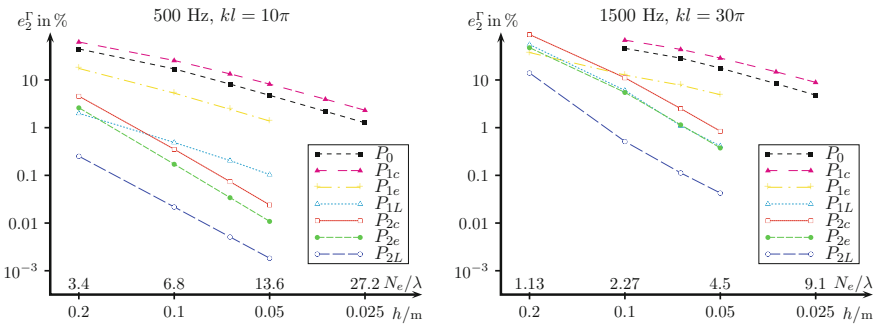


Fig. 11 Long duct, surface error in Euclidean norm, error in terms of element size

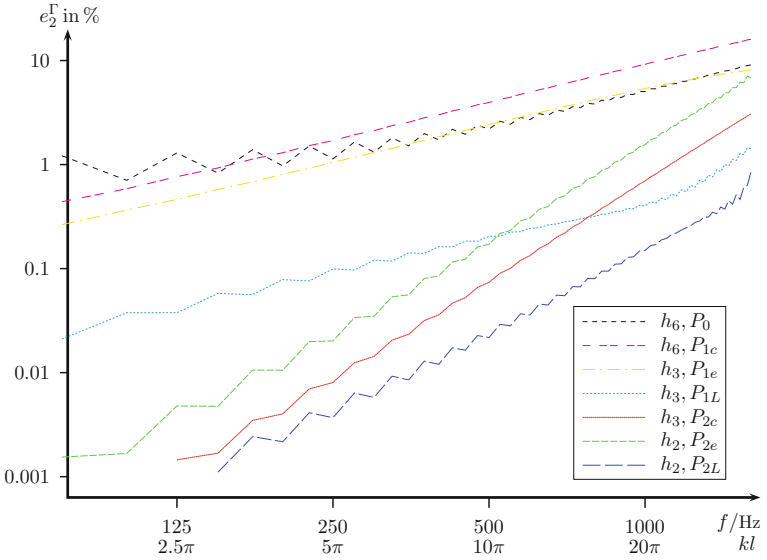


Fig. 12 Long duct, surface error in Euclidean norm: comparison of different element types. Note, that all models have the same degree of freedom of approximately 2520

nodes N and, thus, the degree of freedom, is $N \approx N_e$ for P_0 and P_{1c}, $N \approx 4N_e$ for P_{1e}, P_{1L} and P_{2c} and $N = 9N_e$ for P_{2e} and P_{2L}.

Figure 12 presents the numeric error in terms of frequency for all element types considered in this chapter and for approximately the same degree of freedom. (The models which consist of continuous elements have a degree of freedom of 2522.) Still here, it is obvious that the discontinuous quadratic elements P_{2L} provide the highest accuracy, hence they appear to be the most efficient. However, continuous quadratic elements perform quite well and are more efficient than the discontinuous quadratic elements P_{2e} which differ from the P_{2L} elements only by the location of the collocation points on the element. The performance of linear discontinuous elements P_{1L} is amazing because they are not just much more accurate than the other two linear elements but they even perform better than P_{2e} and P_{2c} in the higher frequency range where between three to six of these elements per wavelength are used. As it has been shown in Marburg (2008), these remarkable results are not observed that clearly when the error is measured in the maximum norm. Then, the performance of elements P_{2L} is very close to that of the continuous quadratic elements P_{2c}. Furthermore, the error curve of the P_{1L} is not as favorable as for the case that the error is measured in the Euclidean norm.

Another remarkable observation consists in the fact that, apart from the very low frequency range, constant elements perform better than linear continuous elements. This is remarkable insofar that linear continuous elements account for a very popular choice of elements in the boundary element collocation method.

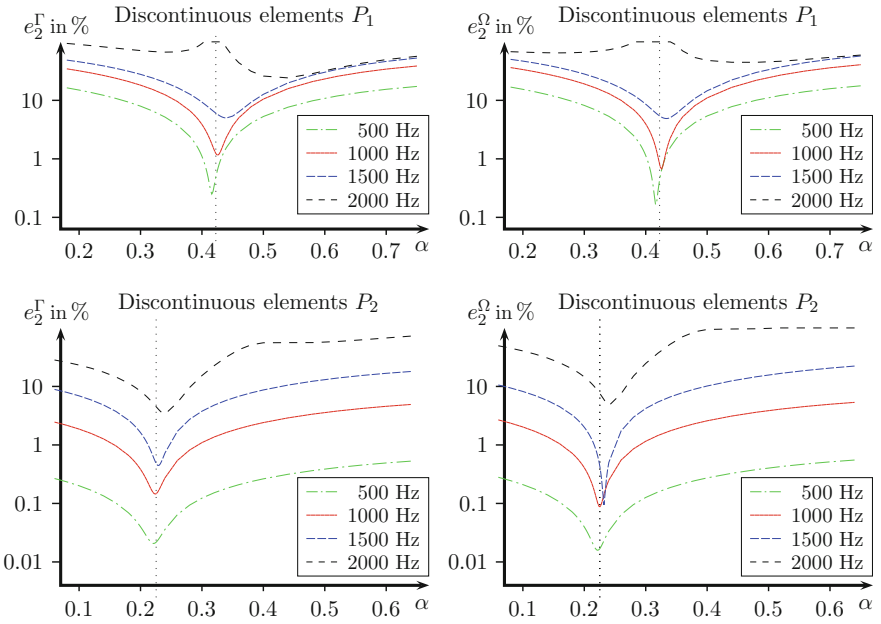


Fig. 13 Long duct, surface error of discontinuous elements in terms of the position of the collocation points on the element

Returning to the discontinuous linear and quadratic elements, the location of the nodal points on the element is investigated. It became obvious in the previous investigations that location of nodes at the zeros of Legendre polynomials provides lower errors compared to an equidistant distribution of nodes on the surface. More general, nodes should be located at zeroes of orthogonal functions being defined on the standard interval $[-1, 1]$. Legendre polynomials account for the simplest selection of orthogonal functions since they are well-known and particularly designed for the interval $[-1, 1]$. Legendre polynomials are orthogonal to themselves or, as it is often written, with respect to the constant and unit weighting function. In what follows, it will be investigated whether or not the zeros of the Legendre polynomials actually account for an optimal position of nodes.

Figure 13 shows the errors e_2^Γ and e_2^Ω in terms of α. The test model h_2 consists of 280 elements. The lowest error is expected at $\alpha = 0.4226$ for linear elements and at $\alpha = 0.2254$ for quadratic elements. Although not exactly fulfilled, it can be seen that an optimal location of nodal values is very close to the zeros of the Legendre polynomials. The optimal value varies with frequency. For low frequencies and low error, lower values of α account for an optimal position. For higher frequencies and, consequently, higher errors an α greater than the zeros of the Legendre polynomials is required for optimal elements. In between, a large frequency range is observed where nodal points are optimally placed as predicted, cf. Atkinson (1997).

Actually, the optimal location of nodes at the zeros of Legendre polynomials refers to pure Neumann problems using the double layer potential operator. Herein, a mixed problem is considered because a Robin boundary condition is applied at one end of the duct. Apparently, the choice of nodes at the zeros of the Legendre polynomials is a good approximation of the optimal location. In case of other operators, i.e. the hypersingular operator as used in the subsequent section, and other boundary conditions, the optimal position of nodes may differ from the one identified here.

It shall be mentioned at this point that for certain frequencies, extremely low errors are gained for field points compared to surface error. The most remarkable example is found for quadratic elements at 1500 Hz. However, significant differences can be found at 500 and 1000 Hz for both, linear and quadratic elements. The error of the solution at the surface and at internal points is almost the same for very low and for higher frequencies.

A similar analysis is, of course, possible for triangular elements. This has been carried out in the previous work of the author, cf. Marburg and Schneider (2003b), Marburg (2008).

However, as a conclusion for this subsection, it shall be emphasized that linear continuous elements are likely to be the worst element choice for boundary element collocation methods. The position of nodal points on the element can influence the accuracy of the solution by one to two orders of magnitude.

Sedan cabin compartment: This example is chosen to examine an irregular mesh which is the result of an automatic mesh generation. Four meshes are investigated. The meshes consist of quadrilaterals and triangles. Their detailed data are given in Marburg (2008). Figure 14 shows the finest mesh. The elements of this mesh have an edge length less than 5 cm.

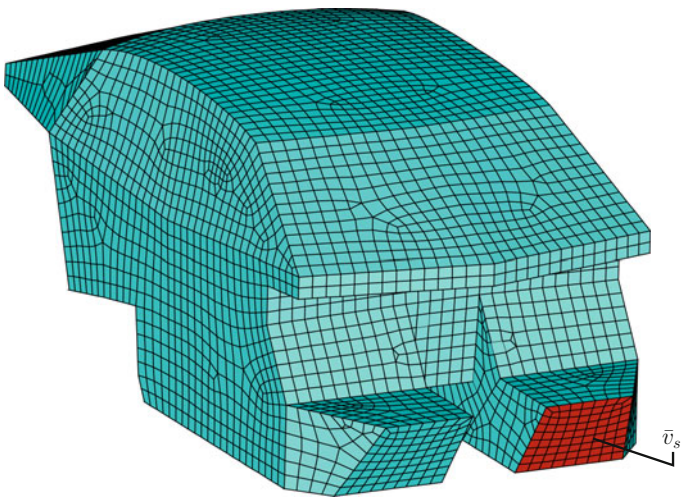


Fig. 14 Finest mesh of sedan cabin compartment, mesh size such that no element length is greater than 5 cm

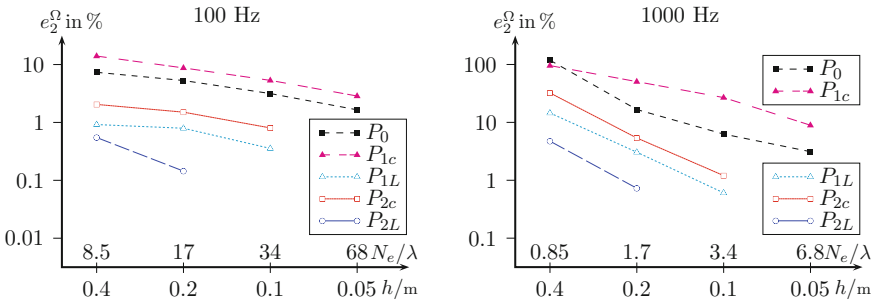


Fig. 15 Sedan cabin compartment, field point error in Euclidean norm, error in terms of element size

A fictitious excitation with uniform normal particle velocity of $\bar{v}_s = 1$ mm/s is applied at the lower left front area. A uniform boundary admittance of

$$Y = \frac{1}{\rho c} \frac{f}{f_0} \qquad (f_0 = 2800 \text{ Hz}) \tag{128}$$

is applied to simulate the absorbing behaviour of the surfaces inside the cabin, cf. Marburg and Hardtke (1999). This value corresponds to experimental measurements of the reverberation time and a corresponding average absorption coefficient. The sound pressure is computed at ten points inside the cabin.

It shall be mentioned that the author is aware that realistic calculations of cabin noise problems are done for frequencies up to (max.) 150. . .200 Hz. However, the major uncertainty of these calculations are structural transfer functions and realistic distributions of the boundary admittance values. It will be shown that even a coarse boundary element mesh for the fluid can give an excellent approximation of the sound pressure field over the entire frequency range.

The reference solution is computed by using discontinuous quadratic elements of size $h \leq 0.1$ m. The associated system of equations has 15744 unknowns. In what follows, we will call this solution our reference solution and all errors are evaluated with respect to this reference.

Looking at the error at internal points in terms of element size, Fig. 15 shows the error functions for different types of elements and different frequencies. The comparison of different element types confirms excellent performance of discontinuous elements. So, it is realized that, in this example again, constant elements give lower error than continuous linear elements. Furthermore, and again, discontinuous linear elements give lower error than continuous quadratic elements. At low frequencies, it is shown that even many elements per wavelength are hardly able to decrease the error below 1%. At higher frequencies with fewer elements per wavelength, the results indicate that the common rules of discretization may be reasonable. However, the author wishes to remark that the error has been determined based on the field point solution at ten points and is based on the Euclidean norm. More points or/and

another norm may have an effect on this result. Nevertheless, the statement about linear continuous elements and discontinuous elements in general should remain valid, independent of the shortcomings of the current approach.

6.6 Conclusion on Choice of Elements

Concluding on the choice of boundary elements for collocation, it has become clear that linear continuous elements should not be used. If a simple element formulation is desired, constant elements have a good performance which is even better than the one for linear elements. However, higher accuracy is only possible to achieve if higher order elements, in particular discontinuous elements are used. It has been shown that discontinuous elements with their nodes at the zeros of orthogonal functions perform the best and may even provide certain superconvergence effects.

7 Irregular Frequencies for Exterior Problems

7.1 The Non-uniqueness Problem

The boundary element formulations as discussed in the previous sections of this chapter are valid for both, interior and exterior problems. If they are applied to the exterior problem, it becomes obvious that the solution is not correct at some frequencies. This is easily demonstrated using a very simple boundary element model. For this, the pulsating sphere is considered. The entire sphere is discretized into 24 elements, i.e. three per octant. These quadrilateral elements approximate the geometry by using quadratic polynomials. We use the data $R = 1$ m, $\rho = 1.3$ kg/m^3, and $c = 340$ m/s, for sphere radius, fluid density, and speed of sound, respectively. A unit particle velocity $v_f (= v_s)$ independent of the location on the surface and frequency is applied. We consider the frequency range up to 2000 Hz. This is equivalent to the relative wavenumber $kR \approx 11.75\pi$. Since the integrand is a highly oscillatory function for high frequencies, we need a much higher order of integration than the one discussed previously. Here, all integrals were evaluated using Gauss-Legendre quadrature rule with 30 integration points per direction, i.e. 900 points per element. This is only possible because the solution for the sound pressure on the surface of the sphere is not oscillating at all. A constant sound pressure on the surface is expected and this is the best approximated by a few constant elements. Insofar, this example is unusual. The analytic expression for the sound pressure magnitude at the surface can be written as

$$\bar{p}(R) \;=\; \rho\, c\, v_f \, \frac{k\, R}{\sqrt{1 \;+\; k^2\, R^2}}, \tag{129}$$

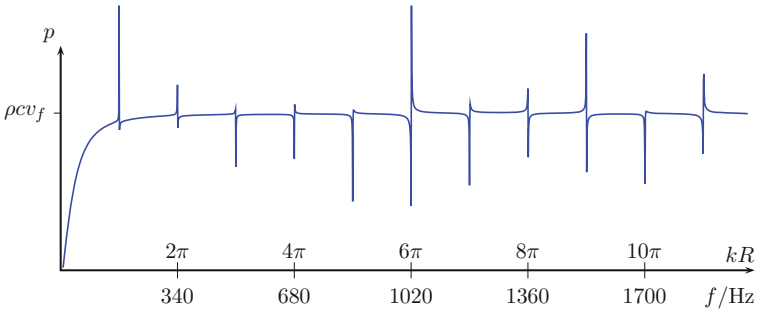


Fig. 16 BEM solution for radiating sphere with uniform surface particle velocity, sound pressure at the surface

where v_f stands for the uniform particle velocity at the surface of the sphere.

In Fig. 16, the numerical solution for the surface sound pressure is given. This surface value of the sound pressure is actually an average over all the nodes. However, as expected, we observed very little difference in the nodal values. The solution for the sound pressure at the surface is well approximated with the exception of a few small frequency regions where the solution obviously fails. These frequencies are often denoted as irregular frequencies and the problem is known as the non-uniqueness problem or the non-uniqueness difficulty.

The non-uniqueness difficulty is not easily explained from the physical ground. Although the exterior domain and the corresponding (imaginary) interior domain share the same boundary, the boundary integral equations for the exterior and the interior problems are still slightly different in two aspects

1. their normal directions are opposite to each other, and
2. their solid angles and, thus, the integral-free terms are different at corners and edges.

The latter is not relevant for discontinuous elements while the former changes the sign of the integral term in Eq. (26). Matrix $\boldsymbol{G}$ is actually the same for interior and exterior problems.

Although the integral equation being very similar, it is hard to just directly compare the interior and exterior boundary integral equations to explain why the non-uniqueness difficulty would occur. Advanced mathematical explanations have been presented in papers half a century ago, cf. Kupradze (1956) and Weyl (1952). A simple mathematical explanation of this phenomenon can be found in Wu and Seybert (1991) and later in Wu (2000a) as well. Since matrix $\boldsymbol{G}$ is the same for interior and exterior domains, it will always become (nearly) singular at eigenfrequencies of the interior Dirichlet problem with homogeneous boundary conditions, i.e. for the problem when the sound pressure at the entire boundary is prescribed to be zero. Such a singularity is also occurring for the matrix $\boldsymbol{H}$ at these eigenfrequencies of the interior Dirichlet problem. With the singularities at both sides of the integral equation (15), these irregular frequencies can be understood as a gap in the sound pressure solu-

tion which only happens at some discrete frequencies. These gaps are not physical at all but occur as a mathematical artefact. While these gaps are infinitely small in the analytical solution of the integral equation, they result in an ill-conditioning of the system matrix in the numerical formulation where the result looks similar to a resonance. Therefore, they are also called spurious modes which are completely unphysical.

It has been shown that regardless of the type of boundary conditions prescribed for the exterior problem (Neumann, Dirichlet, or impedance), see Wu (2000a), the Kirchhoff–Helmholtz integral equation will always fail to yield a unique solution at the eigenfrequencies of the corresponding interior Dirichlet problem. In real-world applications, knowing the exact locations of the irregular frequencies is actually not that important because it is impractical to solve an interior Dirichlet problem first just to find the eigenfrequencies. A more reasonable approach is to always apply some kind of treatment at every frequency to prevent the non-uniqueness from happening. Actually, at high frequencies, the eigenfrequencies are so closely spaced that it is impossible to distinguish the regular frequencies from the irregular frequencies. This is demonstrated in another example of a sphere, which is now a spherical scatterer in a field of plane waves where the numerical solution can be compared to an analytical solution as well.

Again, the sphere is assumed to be rigid, i.e. $Y = 0$, and material data of air are used. The analytical solution for the total sound pressure p is well-known as sum of incident and scattered sound pressures, p_i and p_s, respectively, see for example Ihlenburg (1998)

$$p(r, \vartheta) = p_i + p_s = p_0 \left\{ \mathrm{e}^{\mathrm{i}kr\cos\vartheta} + \sum_{n=0}^{\infty} \mathrm{i}^n (2n+1) \frac{j_n'(kR)}{h_n'(kR)} P_n(\cos\vartheta) h_n(kr) \right\}. \tag{130}$$

In Eq. (130), p_0 represents the sound pressure amplitude of the incident wave; j_n and h_n are the spherical Bessel and Hankel functions of the first kind, respectively. P_n denotes the Legendre polynomial of order n and $R = 1\,\mathrm{m}$ is the radius of the spherical scatterer. Since the problem is axisymmetric, the two parameters r and ϑ allow a complete evaluation in space; r is the distance from the center of the sphere and ϑ is the angle such that the shadow zone is located at $\vartheta = 0$ and the illuminated zone at $\vartheta = \pi$.

The numerical model is a model which uses super-parametric boundary elements for which the geometry is, again, approximated by quadratic quadrilateral elements (9 nodes per element) and the sound pressure is approximated by linear discontinuous boundary elements as described in the previous section. It consists of 1536 boundary elements, i.e. 192 per octant. The element length is approximately 0.1 m, i.e. 64 elements along the diameter of the sphere. Selecting a maximum frequency of 1700 Hz for the analysis, i.e. $kR = 10\pi$, results in a mesh for which 3.4 elements per wavelength are counted. According to the results of the previous section, this should be sufficient.

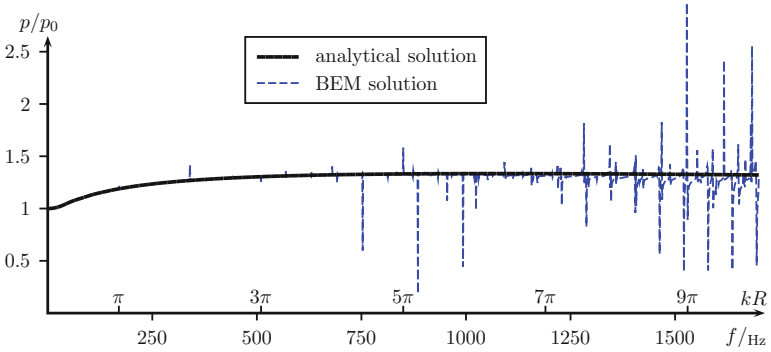


Fig. 17 Spherical scatterer, sound pressure related to magnitude of incident wave at a point $r = 2R, \vartheta = 0$

Figure 17 shows the sound pressure magnitude at $r = 2R$ and $\vartheta = 0$. The comparison confirms that a method to suppress the irregular modes is really necessary. Furthermore, it shows that the requirement for a suitable solution for suppression of these spurious modes is increasing with frequency and the number of irregular frequencies in a certain frequency interval is also increasing with frequency.

The literature knows a number of techniques to suppress these spurious modes. The two most popular techniques, i.e. the Combined Helmholtz Integral Equation Formulation (CHIEF) and the method of Burton and Miller, will be discussed in what follows. Surveys of these and other methods for suppression of spurious modes are given in Wu and Seybert (1991), do Rego Silva (1993), Wu (2000b), Marburg and Amini (2005), Marburg and Wu (2008).

7.2 Combined Helmholtz Integral Equation Formulation

A simple and very popular method to overcome the ill-conditioning of the system matrix $\boldsymbol{H}$ in Eq. (26) consists in an addition of collocation points in Ω_c. This was first proposed in Schenck (1968). This method is known as Combined Helmholtz Integral Equation Formulations or short CHIEF.

In the CHIEF, the system of equations is first set up as given in Eq. (28). Then, additional equations of the type of Eq. (24) are added for the collocation point $\vec{z}_l \in \Omega_c$. In such a case of the location of the collocation point, the integral free term is zero since $c(\vec{z}_l) = 0$, Eq. (14). The additional collocation points are usually referred to as CHIEF points. The system matrix $\boldsymbol{H}$ corresponding to the CHIEF method takes the form

$$\boldsymbol{H} \;=\; \left[\begin{array}{cccc} c_1+\bar{h}_{11} & \bar{h}_{12} & \cdots & \bar{h}_{1n} \\ \bar{h}_{21} & c_2+\bar{h}_{22} & \cdots & \bar{h}_{2n} \\ \vdots & \vdots & & \vdots \\ \bar{h}_{n1} & \bar{h}_{n2} & \cdots & c_n+\bar{h}_{nn} \\ \hline \bar{h}_{n+1\,1} & \bar{h}_{n+1\,2} & \cdots & \bar{h}_{n+1\,n} \\ \vdots & \vdots & & \vdots \\ \bar{h}_{n+m\,1} & \bar{h}_{n+m\,2} & \cdots & \bar{h}_{n+m\,n} \end{array}\right] . \tag{131}$$

The entries $\bar{h}_{lj}$ have been given in Eq. (26). The rectangular system matrix reflects the fact that we now have an over-determined linear system of algebraic equations for the n-dimensional vector of unknowns, $\boldsymbol{p}$, because the additional collocation point in Ω_c does not add an additional unknown sound pressure value. Matrices $\boldsymbol{D}$ and $\boldsymbol{G}$ become rectangular too.

This system is solved in a least squares sense, where the unknown solution is formally given by

$$\boldsymbol{p} \;=\; \left[(\boldsymbol{H}-\boldsymbol{D})^H\,(\boldsymbol{H}-\boldsymbol{D})\right]^{-1}\,(\boldsymbol{H}-\boldsymbol{D})^H\,\boldsymbol{G}\,\boldsymbol{v}_s, \tag{132}$$

with superscript ^{H} denoting Hermitian, i.e. conjugate complex transpose of a matrix. Note that the matrix to be inverted, i.e. $(\boldsymbol{H}-\boldsymbol{D})^H\,(\boldsymbol{H}-\boldsymbol{D})$ is Hermitian. Therefore, a simple conjugate gradient method could be applied. Admittance boundary conditions and sources can be considered in the same way as described in previous sections.

It is well known that when a CHIEF point falls on any of the internal nodal surfaces of the corresponding interior Dirichlet problem, that particular CHIEF point will not provide any constraint effect because sound pressure on any internal nodal surface is automatically zero by definition. For a general radiation/scattering problem, it is unlikely to know the exact locations of the internal nodal surfaces unless the corresponding interior Dirichlet problem is solved first. The problem is compounded by the fact that the internal nodal surfaces are clustered together at high frequencies in such a way that it is almost impossible for a CHIEF point not to fall on any nodal surface. For that reason, there are a number of modifications of the CHIEF. For more details and further references, the reader is referred to the discussion in Marburg and Wu (2008).

It should be noted that CHIEF and all modified CHIEFs only extend the application range of the boundary element formulation from low frequencies to intermediate frequencies. The rank deficiency of the original BEM matrix (without using any CHIEF) is usually greater than one. It has been found in numerical experiments that a single CHIEF point that does not fall on any nodal surfaces still may not be able to provide enough constraint effect. Therefore, a general applicability of CHIEF to get a safe solution is not given.

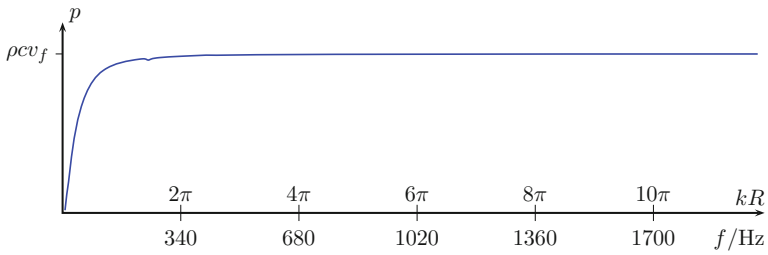


Fig. 18 CHIEF solution using one CHIEF point for radiating sphere with uniform surface particle velocity, sound pressure at the surface

Figure 18 shows the same solution as Fig. 16 with the difference that one CHIEF point in the centre of the sphere has been added. In this simple example, only one CHIEF point is sufficient to suppress all relevant spurious modes. As mentioned above, this cannot be generalized. Usually, the problem is more complex.

An update of Fig. 17, i.e. scattering from a sphere, is given in Fig. 19 for the same boundary element model as above but with different number of CHIEF points. Obviously, the spurious modes are not sufficiently suppressed with only one point (in the sphere's center) for this load case. When using 37 arbitrarily but well distributed CHIEF points, the solution looks much smoother and more accurate. Increasing the number of CHIEF points to 260 is further improving the solution and we can hardly identify any spurious mode in the frequency range under consideration.

7.3 Method of Burton and Miller

An alternative approach to CHIEF was proposed in Burton and Miller (1971). The idea of this technique was originally proposed for the Dirichlet problem in two papers in parallel, i.e. Panič (1965), Brakhage and Werner (1965), see also Kussmaul (1969). All these approaches have in common that they are using the normal derivative of an original integral equation. As the problem results in rather long equations, the author decides that the admittance remains unconsidered in what follows. Hence, $Y = 0$ and, thus, $\boldsymbol{D} = \boldsymbol{0}$. Taking the normal derivative of the original integral equation means that in our case, the normal derivative of the Kirchhoff–Helmholtz integral equation (15) is used. For the further use, Eq. (15) is rewritten as

$$\Psi_1(\vec{\boldsymbol{y}}) = c(\vec{\boldsymbol{y}})p(\vec{\boldsymbol{y}}) + \int_\Gamma \frac{\partial G(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}})}{\partial n(\vec{\boldsymbol{x}})} p(\vec{\boldsymbol{x}}) d\Gamma(\vec{\boldsymbol{x}}) - \int_\Gamma G(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}}) a v_f(\vec{\boldsymbol{x}}) d\Gamma(\vec{\boldsymbol{x}}) = 0. \tag{133}$$

This integral equation is often referred to as the first integral equation. The normal derivative at the surface point $\vec{\boldsymbol{y}}$ is yielded as

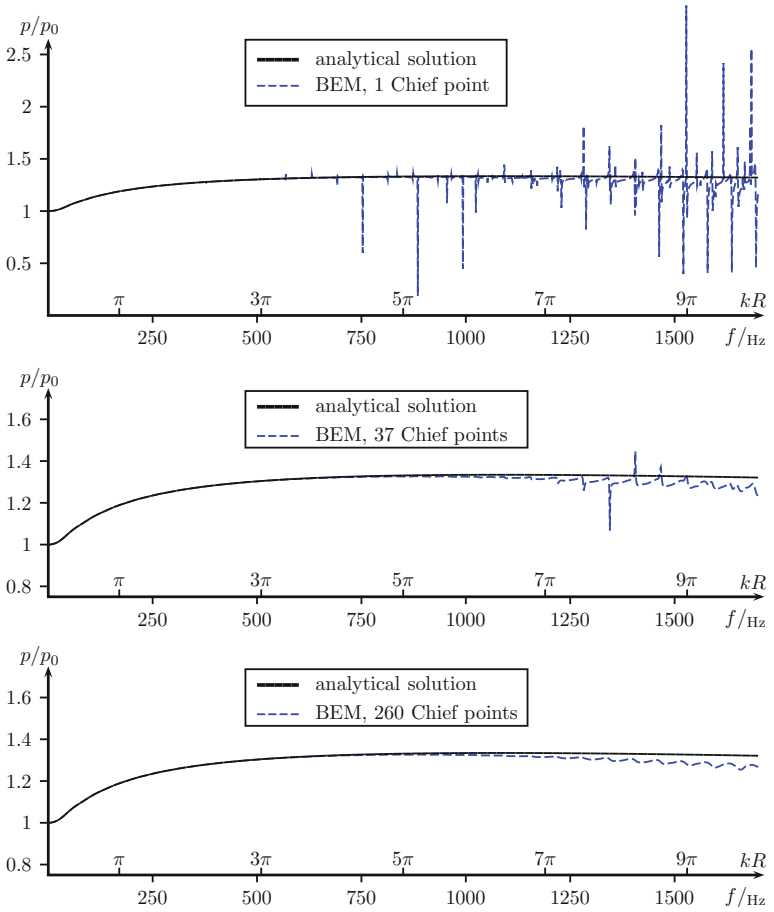


Fig. 19 Spherical scatterer, sound pressure related to magnitude of incident wave at a point $r = 2R$, $\vartheta = 0$ for different number of CHIEF points

$$\Psi_2(\vec{\mathbf{y}}) = \frac{\partial \Psi_1(\vec{\mathbf{y}})}{\partial n(\vec{\mathbf{y}})} = 0 = \tag{134}$$

$$= c(\vec{\mathbf{y}}) \frac{\partial p(\vec{\mathbf{y}})}{\partial n(\vec{\mathbf{y}})} + \int_\Gamma \frac{\partial^2 G(\vec{\mathbf{x}}, \vec{\mathbf{y}})}{\partial n(\vec{\mathbf{x}}) \partial n(\vec{\mathbf{y}})} \, p(\vec{\mathbf{x}}) d\Gamma(\vec{\mathbf{x}}) - sk \int_\Gamma \frac{\partial G(\vec{\mathbf{x}}, \vec{\mathbf{y}})}{\partial n(\vec{\mathbf{y}})} \, v_s(\vec{\mathbf{x}}) d\Gamma(\vec{\mathbf{x}})$$

and is often referred to as the second integral equation. It consists of another integral free term and the hypersingular operator whereas the second integral is known as the adjoint double layer potential.

Discretization by collocation of the second integral equation leads to the matrices $\boldsymbol{F}$ for the integral free term added to the adjoint double layer potential and $\boldsymbol{E}$ for the hypersingular operator with the entries determined as

$$f_{lj} = -s\,k\,c(\vec{z}_l)\,\delta_{lj} \;+\; s\,k \int_\Gamma \frac{\partial G(\vec{x},\vec{z}_l)}{\partial n(\vec{z}_l)} \phi_j(\vec{x})\,d\Gamma(\vec{x}) \tag{135}$$

and

$$e_{lj} = \int_\Gamma \frac{\partial^2 G(\vec{x},\vec{z}_l)}{\partial n(\vec{x})\partial n(\vec{z}_l)}\,\phi_j(\vec{x})\,d\Gamma(\vec{x})\;. \tag{136}$$

The second derivative of the Green's function is given as

$$\frac{\partial^2 G(\vec{x},\vec{y})}{\partial n(\vec{x})\partial n(\vec{y})} = \frac{1}{4\pi r^3}\left[(3 - 3\mathrm{i}kr - k^2r^2)\frac{\partial r}{\partial n(\vec{x})}\frac{\partial r}{\partial n(\vec{y})} + \right. \\ \left. + (1 - \mathrm{i}kr)\;\vec{n}(\vec{x})\cdot\vec{n}(\vec{y})\right] e^{\,\mathrm{i}kr}. \tag{137}$$

Note that in this second derivative of the Green's function, normal vectors are neither appearing alone nor as single normal derivatives. Only scalar products of two normal vectors or products of normal derivatives are yielded. This means that matrix $\boldsymbol{E}$ is exactly the same for interior and exterior domains. Discretizing the second integral equation and formulating matrices $\boldsymbol{E}$ and $\boldsymbol{F}$ as given in Eqs. (136) and (135), respectively, we get

$$\boldsymbol{E}\,\boldsymbol{p} = \boldsymbol{F}\,\boldsymbol{v}_s\,. \tag{138}$$

Similar to the discussion on the determination of the integral free term, the sum of row elements for matrix $\boldsymbol{E}$ at the frequency of zero must be zero. This knowledge can be used for regularization of the hypersingular integral in Eq. (136). If this is not used for evaluation of matrix $\boldsymbol{E}$, this knowledge can, at least, be used for checking the accuracy of the integration procedure.

The hypersingular operator is a pseudo-differential operator of order +1; meaning that it behaves essentially as a first order differential operator. Because of this, in Eq. (134), the unknown function $p(\vec{x})$ requires higher continuity than that for Eq. (133). More specifically, the hypersingular operator requires C^1 continuity of the function at collocation points. As discussed already in the previous section, it is quite common and indeed advantageous to use discontinuous boundary elements since they fulfill this condition. Note that they even have C^∞ continuity at collocation points. Alternatively, it is possible to use Galerkin discretization and continuous elements.

The method of Burton and Miller uses a linear combination of the first and the second integral equation such that

$$\Psi_1(\vec{y}) + \eta\Psi_2(\vec{y}) = 0\,, \tag{139}$$

where the coupling parameter η is introduced. There are not many requirements for choosing this parameter. The main requirement consists in the fact, that the

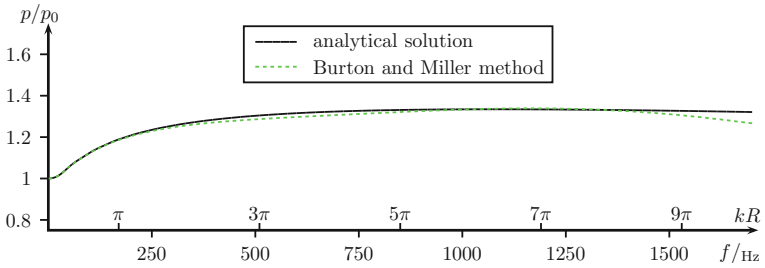


Fig. 20 Spherical scatterer, sound pressure related to magnitude of incident wave at a point $r = 2R$, $\vartheta = 0$ for Burton and Miller method

parameter must not be real but complex. The non-zero imaginary part guarantees a unique solution, at least analytically. Numerically, it is sensible to choose the coupling parameter purely imaginary (Meyer et al. 1978). While the literature mostly suggested $\eta = \mathrm{i}/k$ as the optimal value for higher frequencies, it is worth mentioning that in some formulations and including the one presented here, it should be negative, i.e. $\eta = -\mathrm{i}/k$, see Marburg (2016a). Note, that as it was shown in that paper, there is a substantial amount of literature proposing a wrong sign of the coupling parameter. The reason why the coupling parameter with wrong sign can have a significant effect on the solution at all is currently a topic of research in mathematics, see for example Galkowski et al. (2016).

Applying Eq. (139) to the discretized formulations, it can be written as

$$[\boldsymbol{H} + \eta \boldsymbol{E}]\ \boldsymbol{p} = [\boldsymbol{G} + \eta \boldsymbol{F}]\ \boldsymbol{v}_s \tag{140}$$

which is solved for $\boldsymbol{p}$. Admittance boundary conditions and sources can be considered in the same way as described in previous sections.

Although setting up four matrices or, in the presence of admittance boundary conditions, even five, it is easily possible to organize a computer code such that while the matrices are evaluated all at once, only one system matrix is stored in memory. Therefore, the memory requirements of the Burton and Miller method are negligibly higher than for the ordinary BEM solution. Also, the complexity of the Burton and Miller method is hardly higher than for the solution of the Kirchhoff–Helmholtz integral equation.

Figure 20 shows the results of the Burton and Miller method for the same scattering example as in the two previous subsections in Figs. 17 and 19. The boundary element model remains the same but, of course, without any CHIEF points. It is clear, that the Burton and Miller method is able to efficiently suppress the spurious modes. A recent paper, cf. Zheng et al. (2015), has shown that the Burton and Miller method shifts the unwanted eigenvalues far into the complex plane such that they are efficiently damped out. Consequently, they are still there but remain virtually invisible.

Figure 21 compares the analytical solution with the solution by CHIEF with 260 CHIEF points and the Burton and Miller method for the scattering problem but using

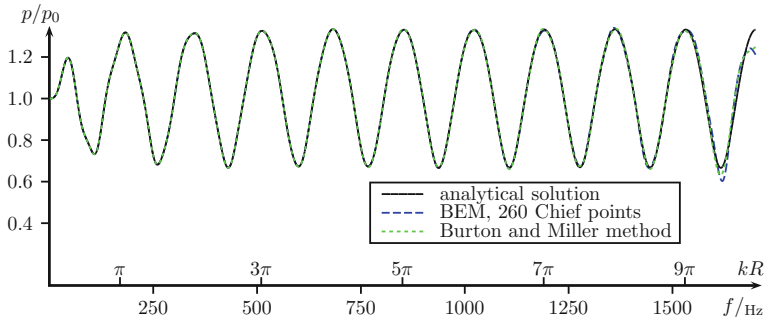


Fig. 21 Spherical scatterer, sound pressure related to magnitude of incident wave at a point $r = 2R$, $\vartheta = \pi$, comparison between analytical solution, solution for 260 Chief points and for the Burton and Miller method

another location. Again, the three solutions are pretty close to each other and only close to the highest frequency, some differences can be recognized. Apparently, the validity of the boundary element model is limited in that region.

A second example considers the sound radiation from a diesel engine, which is taken from the author's papers (Marburg and Wu 2008; Fritze et al. 2009). For that, the sound power solutions of the Burton and Miller method, the lumped parameter model, i.e. the sound power estimation based on the Rayleigh integral, and the equivalent radiated power approximation, i.e. sound power estimation based on the assumption that $p(\vec{\boldsymbol{x}}) = \rho_0 c v_s(\vec{\boldsymbol{x}})$ for $\vec{\boldsymbol{x}} \in \Gamma$, are compared. The fluid surface model contains 20172 nodes and 21497 constant elements. The problem is solved for the frequency

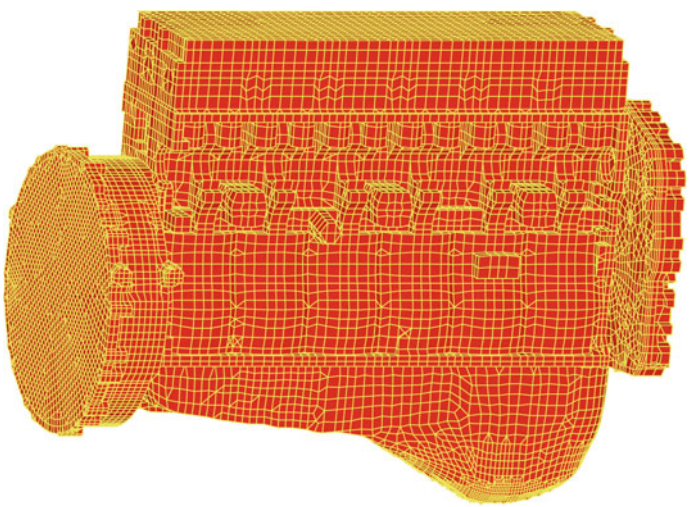


Fig. 22 Boundary element model of a six cylinder diesel engine

range up to 3000 Hz. The boundary element model is presented in Fig. 22. The realistic excitation of the acoustic field is applied by defining the particle velocity over the surface at each investigated frequency.

The particle velocity distribution over the engine's surface for a certain operation condition was computed and provided by the AVL/ACC Graz (Austria). Originally, the particle velocity was given on the mesh of linear continuous elements. The

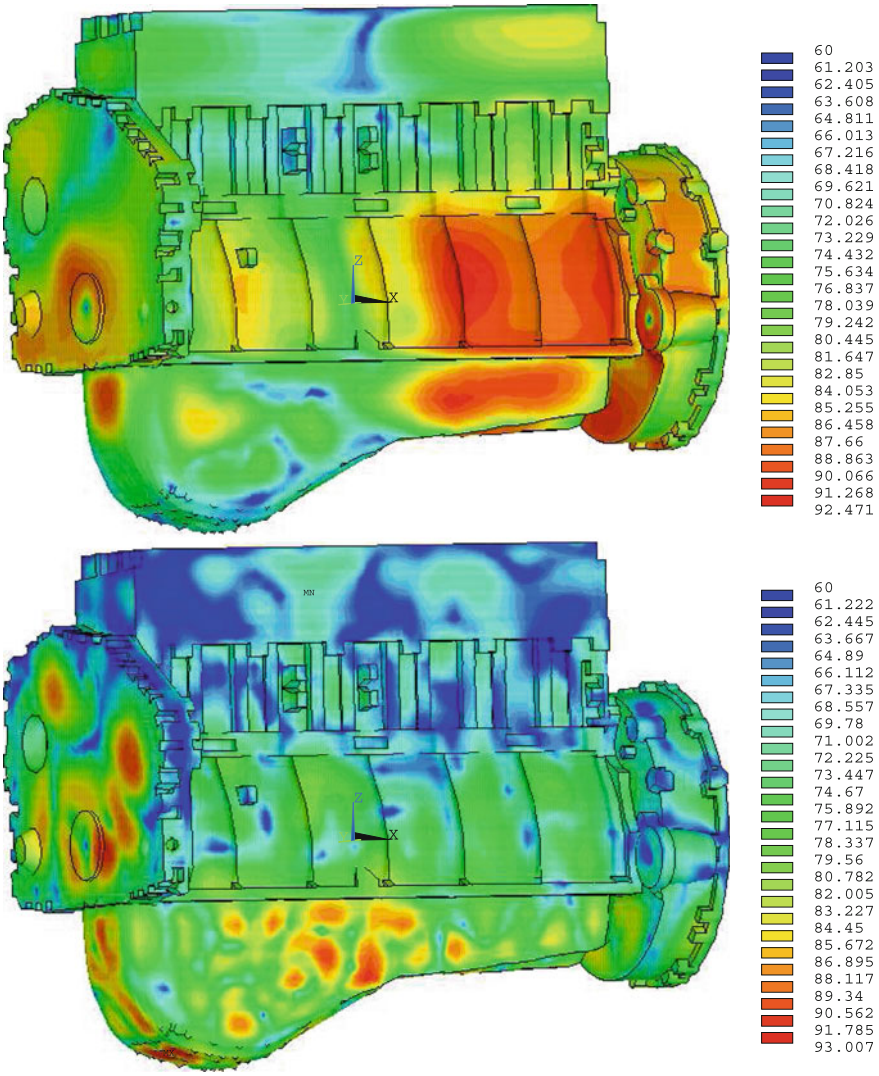


Fig. 23 Acoustic radiation from diesel engine, sound intensity levels at 500 Hz (*top*) and 2200 Hz (*bottom*)

piecewise constant particle velocity data which is used for our simulations can be understood as an average of the normal velocity on each element.

To provide the reader with a vivid impression of a solution, the sound intensity levels at the engine's surface are visualized for two specified frequencies, i.e. approximately 500 Hz and approximately 2200 Hz, cf. Fig. 23. The sound intensity vector $\vec{\boldsymbol{I}}$ is evaluated as

$$\vec{\boldsymbol{I}}(\vec{\boldsymbol{x}}) = \frac{1}{2}\Re\left\{p(\vec{\boldsymbol{x}})\vec{\boldsymbol{v}}_f^*(\vec{\boldsymbol{x}})\right\} \qquad \text{with} \quad \vec{\boldsymbol{x}} \in \Omega \cup \Gamma\,, \tag{141}$$

whereas the surface intensity I_n is given (for rigid boundary conditions) as

$$I_n(\vec{\boldsymbol{x}}) = \frac{1}{2}\Re\left\{p(\vec{\boldsymbol{x}})v_s^*(\vec{\boldsymbol{x}})\right\} \qquad \text{with} \quad \vec{\boldsymbol{x}} \in \Gamma_s\,. \tag{142}$$

Therein, the arbitrary surface Γ_s can also be Γ, very similar to the discussion in the context of Eqs. (86) and (87) where the radiated sound power had been introduced. The sound intensity level gives a localized impression of the regions where most contributions to the radiated sound power are generated. As shown in Fritze et al. (2009), these regions are not matching well with the particle velocity at 500 Hz but much better at the higher frequency range such as for 2200 Hz. At the lower frequency, the velocity pattern is much finer structured than the sound pressure and the intensity levels. At the higher frequencies, these quantities approach each other. At 2200 Hz, the surface plots of particle velocity and sound pressure look more similar in such a way that the sound pressure level distribution appears almost as detailed as the distribution of the particle velocity. So does the sound intensity as shown in Fig. 23. Clearly, the boundary element mesh is even too fine for the highest frequency of 3000 Hz. However, problems introduced by the coupling between the structure and the fluid mesh are avoided, since the outer surface of the structural mesh is directly used as the fluid BE mesh.

It is a common technique to determine the radiation directivity by evaluating the sound intensity either on certain planar panels or on an enveloping spherical surface. Both are shown in Fig. 24. Of course, the intensity levels at the planar panel rather close to the engine are much higher than those on the enveloping surface. However, the reader may get at least an idea what these intensity plots may be useful for.

The evaluation of the intensity vector on Eq. (141) requires the knowledge of the particle velocity vector. Similar to the determination of the hypersingular boundary integral equation (134), the particle velocity can be directly evaluated by differentiating the representation formula (75) as

$$\begin{aligned}\vec{\boldsymbol{v}}_f(\vec{\boldsymbol{y}}) &= \frac{1}{sk}\frac{\partial p(\vec{\boldsymbol{y}})}{\partial \vec{\boldsymbol{\nu}}} \\ &= \int_\Gamma \frac{\partial G(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}})}{\partial \vec{\boldsymbol{\nu}}}\, v_f(\vec{\boldsymbol{x}})d\Gamma(\vec{\boldsymbol{x}}) - \frac{1}{sk}\int_\Gamma \frac{\partial^2 G(\vec{\boldsymbol{x}},\vec{\boldsymbol{y}})}{\partial n(\vec{\boldsymbol{x}})\partial \vec{\boldsymbol{\nu}}}\, p(\vec{\boldsymbol{x}})d\Gamma(\vec{\boldsymbol{x}})\,,\end{aligned} \tag{143}$$

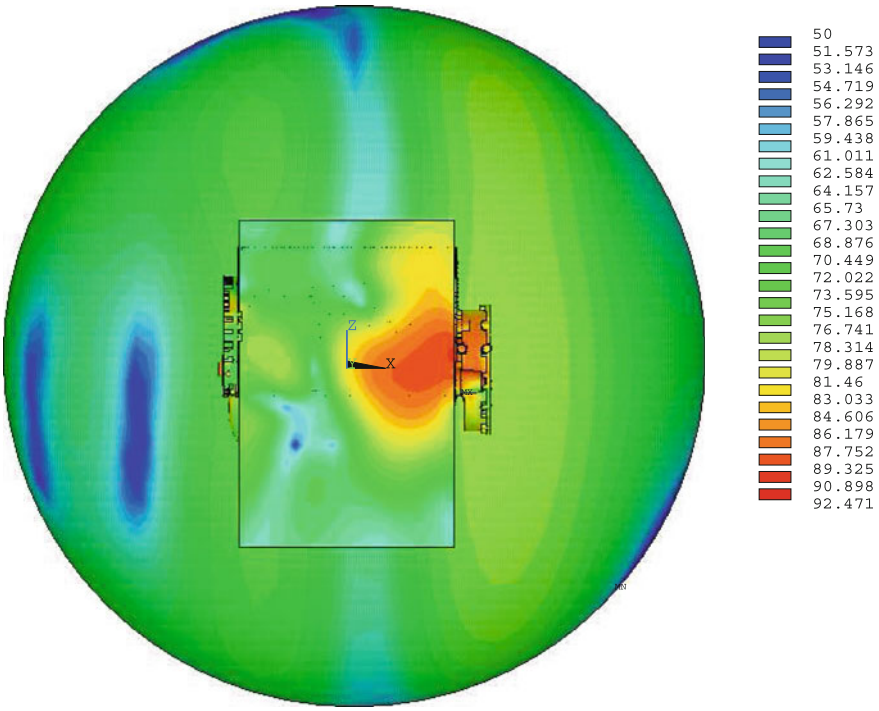


Fig. 24 Acoustic radiation from diesel engine, sound intensity levels at 500 Hz at planar panel and on enveloping surface

where $\vec{\nu}$ is the normal vector on either the planar area or the enveloping sphere. It is one of the advantages of the boundary element technique that the derivative of field point quantities, e.g. the particle velocity vector $\vec{v}_f$ in this case, can be determined analytically. This is usually much more accurate than numerical differentiation by a finite difference approach.

The radiated sound power evaluated based on the Burton and Miller solution and Eq. (140) is compared with the Equivalent Radiated Power (ERP) for which the surface sound pressure is directly and locally determined based on simple assumption of equivalence to the particle velocity, see Fig. 25. (ERP assumes radiation efficiency of 1.) There, the ERP values agree with the BEM solution the better, the higher the frequencies are. Actually, it is quite surprising that the very simple approximation of the ERP catches the behaviour of the Burton and Miller solution with this accuracy. In the lower frequency range, the difference of the approximate solution and the BEM reference are a few decibels. However, peaks and valleys in the curves are found at the same frequencies. It turned out that the LPM solution would be hardly different from the Burton and Miller solution. It is therefore omitted in Fig. 25.

The sound power level differences between the approximate solutions of ERP and LPM are depicted in Fig. 26. Therein, it is confirmed that this difference is only a few

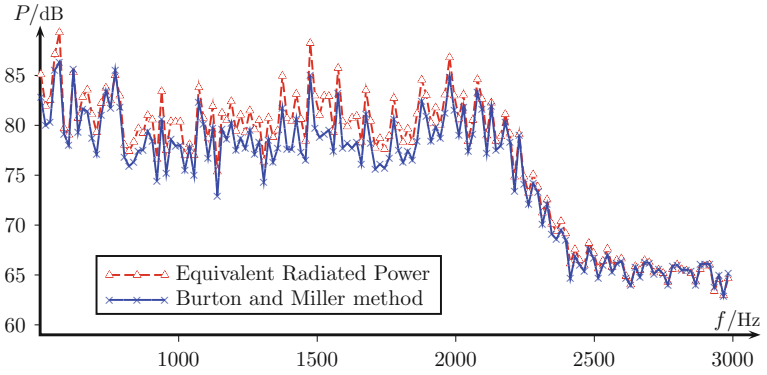


Fig. 25 Acoustic radiation from diesel engine, sound power level over frequency, comparison of BEM solution with the Burton and Miller methods and the approximate solution based on equivalent radiated power

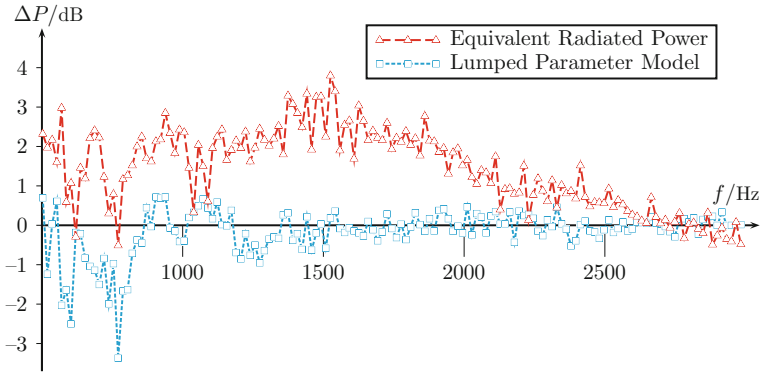


Fig. 26 Acoustic radiation from diesel engine, sound power level deviation from Burton and Miller solution for approximate solution based on equivalent radiated power and the lumped parameter model

decibels for ERP and clearly diminishing in the higher frequency range. It remains below 4 dB over the entire frequency range whereas the actual differences between two neighboring frequency points are up to 8 or 9 dB, not speaking about uncertainties of the model parameters. It is a property of ERP that it usually overestimates the radiated sound power since it assumes a radiation efficiency of 1. However, the example shows that, in some cases, the actual radiation efficiency is greater than 1 and thus, ERP would be underestimating the actual radiated sound power.

The LPM solution is even more accurate than ERP. While there are deviations of a few decibels in the frequency range between 500 and 800 Hz, the LPM solution deviates less than 1 dB in the frequency range beyond 800 Hz.

While the LPM solution does not require the solution of a system of equations, computation of the double sum required for the LPM, see Eq. (16) in Fritze et al.

(2009) is still requiring $O(N_e^2)$ floating point operations, which is asymptotically the same as required for ordinary BEM, CHIEF and the Burton and Miller method. ERP is much faster since it is evaluated based on a single sum, see Eq. (12) in Fritze et al. (2009), and thus requiring only $O(N_e)$ floating point operations. It can be assumed that the number of elements is (asymptotically) of the same order of magnitude as the number of nodes.

8 Fast Boundary Element Techniques

8.1 Computational Costs

Efficiency as well as the variety of numerical methods have been increasing in parallel to the development of digital computers and their increasing processor speed. Often, new methods are developed and compared to existing methods. Unfortunately, results and performance are not always fairly compared between different methods, e.g. the finite element method and the boundary element method. The paper by Harari and Hughes (1992) presented a fair comparison between FEM and BEM for acoustic problems. It started from the idea that both methods require a certain number of elements per wavelength to produce results of a certain, prescribed and equal accuracy. According to this reference level, Harari and Hughes argued that the degree of freedom N for a surface (boundary element) mesh is of order $N = O((kl)^2)$ (k is the wavenumber, l is a representative, e.g. the largest, distance of the model), whereas for a volume mesh it is $N = O((kl)^3)$. The solution procedure for a conventional BEM (with iterative solution of the linear system of equations) requires $O(N^2)$ memory resources and its complexity (number of operations) shows the $O(N^2)$ behavior too. FEM benefits from its sparse matrices for which both memory and complexity show an $O(N)$ behavior. This leads to memory requirements and complexity for BEM of $O((kl)^4)$ and for FEM of $O((kl)^3)$. This shows that conventional boundary element solutions are less efficient than finite element solutions.

Hence, the conventional boundary element solution as described in the previous subsections requires $O(N^2)$ operations keeping in mind that the degree of freedom is $N = O((kl)^2)$. The same holds for memory requirements. The author has run a test example on a modern personal desktop computer to determine numbers for memory requirement and computation time. The example is a scattering cube with 1 m edge length in air with material data as in the previous examples. The scattering problem with an incoming plane wave is solved for the Kirchhoff–Helmholtz integral equations at low frequencies, i.e. below 100 Hz, and in the higher frequency range, i.e. between 3300 and 3400 Hz. The boundary element model consists of 9600 square elements which comes to 40 × 40 elements per face and, thus, an element length of 2.5 cm. With a wavelength of 10 cm at 3400 Hz, this results in 4 elements per wavelength in this high frequency range. The model is using linear discontinuous elements which results in a degree of freedom of 38400. Memory requirements are such that

Table 2 Scattering from a cube (38400 collocation points): Comparison of computation times for the ordinary BEM as the solution of the Kirchhoff–Helmholtz integral equation (K–H IE), CHIEF and the method of Burton and Miller. Results are shown for very low frequencies and a frequency for which 4 discontinuous linear elements per wavelength are used. Setup time and solution time are compared. N_{mvp} gives the average of 10 evaluations at 10 different frequencies

Method equation	Frequency [Hz]	Setup time [s]	Solution time [s]	Matrix-vector products N_{mvp}
K–H IE (66)	$\leq$100	360	54	18
K–H IE (66)	$\approx$3400	360	738	248
CHIEF (132)	$\leq$100	363	199	68
CHIEF (132)	$\approx$3400	363	1386	473
B&M (140)	$\leq$100	391	284	95
B&M (140)	$\approx$3400	391	113	38

the Kirchhoff–Helmholtz solution needs 22 Gigabytes at low frequencies and 22.5 Gigabytes at high frequencies. (The difference is due to the larger subspace of the GMRes solver.) This can be extrapolated for a larger degree of freedom. Hence, for $N = 10^5$, a memory of 149 Gigabytes is required and $N = 2 \cdot 10^5$ demands a memory of 596 Gigabytes, which is, at least at the moment, beyond realistic expectations for a personal computer.

Of course, computation time is always depending on the implementation. The Fortran90 code Akusta which is used herein has been developed in a straightforward manner. As described in one of the previous sections, it performs an outer loop over all collocation points and an inner loop over all elements. The setup of the matrices can be accelerated by exchanging these loops. For a better readability of the code, this has not been done. Integration is performed adaptively as described in this chapter. The system of equations is solved by the GMRes iterative solver without preconditioning. Table 2 shows the computation times for the solution based on the Kirchhoff–Helmholtz integral equation (K–H IE), cf. Eq. (66), the combined Helmholtz integral equation formulation using 153 CHIEF points, cf. Eq. (132), and the Burton and Miller (B&M) method, cf. Eq. (140). As mentioned above, test cases are run at very low frequencies below 100 Hz and at frequencies of approximately 3400 Hz, where the element size is chosen such that it could be used in realistic applications. An arbitrarily chosen residual of 10^{-10} is demanded for the iterative solution of the systems of equation. The number of matrix-vector products N_{mvp} is an average of ten computations at ten different frequencies in the vicinity of the frequency shown.

The setup time is quite similar for the three methods. CHIEF requires marginally more time than the ordinary BEM based on the Kirchhoff–Helmholtz integral equation. This is due to the additional equations which are generated by the 153 CHIEF points in addition to the 38400 collocation points. The Burton and Miller formulation is slightly more expensive and requires less than 10% more computation time than the setup for the other two methods. Solution of the system of equations with an

iterative solver requires repeated matrix vector products. Each of these matrix vector products takes a little less than three seconds. In the low frequency range where there is no treatment of irregular frequencies required, solution time of the ordinary BEM is much less than the setup time. This is similar for CHIEF, which requires more iterations because of the solution of the overdetermined system of equations in a least squares sense for which the condition number is approximately squared compared to the original one. Furthermore, two matrix-vector products are required per iteration. In the very low frequency range where the Burton and Miller formulation with a coupling parameter $-\mathrm{i}/k$ is dominated by the hypersingular operator, even more solution time than for the other two formulations is required. It remains clearly below the setup time though. The situation is different for the higher frequency range since the number of iterations per solution of the system of equations is dramatically increasing for the ordinary BEM and for CHIEF. This is due to a poor condition at high frequencies. The Burton and Miller formulation results in a decreasing number of iterations since the resulting system of equations is very well conditioned, see also results in Marburg (2016a). Therefore, even a larger residual would be sufficient here. When comparing the overall solution time in the higher frequency range, we have 1098, 1749 and 504 s for the ordinary BEM, for CHIEF and for the Burton and Miller formulation, respectively.

Extrapolating these computation times to larger models and basing this on the 504 s of the Burton and Miller solution and the approximately 3 seconds for one matrix-vector product, it can be estimated that for 10^5 collocation points, a solution time of approximately 57 min with a single matrix-vector product taking 20 s becomes necessary. For $2 \cdot 10^5$ collocation points, we arrive at 228 min and 80 s, respectively. For a degree of freedom of 10^6, it is almost 4 days and more than 30 min per matrix-vector product. However, it should be kept in mind, that currently there is no ordinary desktop computer which is able to store such a matrix in memory. Therefore, there is a demand for techniques which are able to solve problems of this size.

8.2 Basic Idea of Fast BEM

It is the idea of fast boundary element techniques to accelerate the solution such that both, complexity and memory requirements, approach $O(N) = O((kl)^2)$. There are a number of these techniques and they all are based on the idea that the matrix-vector product for the iterative solution can be substantially accelerated. The linear system of Eqs. (28), (35), (66), (73), (74), (132), and (140) can always be written in the form

$$\boldsymbol{A}\,\boldsymbol{p} = \boldsymbol{b}. \tag{144}$$

The iterative solution of this system of equations requires the matrix-vector product of the system matrix $\boldsymbol{A}$ and the current approximation of the solution vector which is $\boldsymbol{p}$ in our case. The current approximation is denoted as $\boldsymbol{p}_k$. Since the entries of the system matrix $\boldsymbol{A}$ depend on the distance between the collocation point and the

current element, see (101), it is fully populated. While the exponential function is smooth (but oscillating), the $1/r$ and $1/r^2$ terms become singular or nearly singular if collocation point and element are close. For greater r, these functions are rather smooth. Therefore, it can be reasonable to split up the matrix-vector product into a near-field part of $\boldsymbol{A}$ for short distances and a far-field part for large distances. The fundamental idea of fast boundary element techniques is to approximate the matrix-vector product of the far-field part of $\boldsymbol{A}$ and the current approximation $\boldsymbol{p}_k$ as

$$\boldsymbol{y} = \boldsymbol{A}\,\boldsymbol{p}_k = (\boldsymbol{A}_{\text{near}} + \boldsymbol{A}_{\text{far}})\,\boldsymbol{p}_k = \boldsymbol{A}_{\text{near}}\,\boldsymbol{p}_k + \boldsymbol{y}_{\text{far}}. \tag{145}$$

There, the near-field matrix $\boldsymbol{A}_{\text{near}}$ is sparse and banded. It is evaluated as usual in BEM and the size of the near field depends on the method and on certain other criteria such as accuracy.

The determination of $\boldsymbol{y}_{\text{far}}$ accounts for the specific feature of all the different methods among the fast boundary element techniques. There are some techniques which are quite popular in the context of problems which are not involving waves, e.g. panel clustering and wavelet transformation. The methods are quite efficient and, thus, well applicable if high accuracy is required. This, however, is not the case for practical applications of acoustics where errors of $>1\%$ are allowed and the element size is of the same order of magnitude as the wavelength. Other techniques seem to be well applicable to acoustic problems. These are

- the regular grid method, cf. Bespalov (2000), Schneider (2003),
- multilevel fast multipole method, cf. Schneider (2003), Gumerov and Duraiswami (2004), Sakuma et al. (2008), Gaul et al. (2008), Gumerov and Duraiswami (2009), Liu (2009), Wu et al. (2013), Wilkes and Duncan (2015), and even including shape sensitivity analysis, cf. Chen et al. (2016)
- hierarchical matrices, also known as $\mathcal{H}$ matrices for which adaptive cross approximation is often used, cf. Bebendorf (2008), Brancati et al. (2011), Bebendorf et al. (2015)
- and hybrid methods combining the fast multipole techniques and $\mathcal{H}$ matrices with adaptive cross approximation, cf. Messner et al. (2012), Liu et al. (2017).

The paper by Brunner et al. (2010) compares the performance of a multilevel fast multipole technique and the use of hierarchical matrices by adaptive cross approximation. Although interesting, these results cannot be generalized. Overall, there is no doubt that the fast multipole method is the most popular method among these fast boundary element techniques so far.

The basic idea of the fast multipole method is comparable with a telephone company, see Fig. 27. In the very early times of telephones with no or only very limited infrastructure, it was necessary to have a direct connection between all participants, cf. upper subfigure of Fig. 27. Later on, connections of all participants in a certain region are summarized in a hub H_m while connections of participants in another region are summarized in the hub H_n. Such a strategy can even be applied on several different levels which would result in a multilevel method, see for example Gieberman (2001). This can be understood as phone connections in different countries

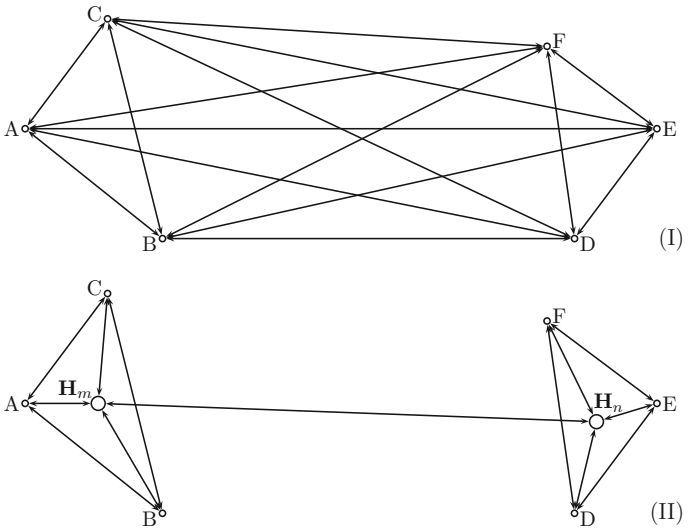


Fig. 27 Basic idea of the fast multipole method: while in the original BEM all collocation points are directly connected as in the early times of the telephone (*I*), a more efficient technique consists in the introduction of hubs in certain regions and only setup connections between these hubs (*II*)

with a hub in a certain district, the next hub in a larger city and another hub as the main connection point in the country whereas this hub corresponds to the major hub in another country and so on. It is obvious that, for a large number of connections, this is most likely the most efficient strategy to connect as many participants as possible.

It is easy to put such a strategy into a mathematical framework. It was mentioned above that the main problem consists in the kernel function of the integration which depends on the distance function $r = |\vec{\boldsymbol{x}} - \vec{\boldsymbol{z}}_l|$. This distance function can be rewritten using the hubs located at $\vec{\boldsymbol{H}}_x$ and $\vec{\boldsymbol{H}}_z$ as

$$\vec{\boldsymbol{r}} = \vec{\boldsymbol{x}} - \vec{\boldsymbol{z}}_l = (\vec{\boldsymbol{x}} - \vec{\boldsymbol{H}}_x) + (\vec{\boldsymbol{H}}_x - \vec{\boldsymbol{H}}_z) + (\vec{\boldsymbol{H}}_z - \vec{\boldsymbol{z}}_l) = \vec{\boldsymbol{r}}_x(\vec{\boldsymbol{x}}) + \vec{\boldsymbol{r}}_H + \vec{\boldsymbol{r}}_z(\vec{\boldsymbol{z}}_l). \tag{146}$$

Although not looking that spectacular, such an approach allows separation of variables such that the kernel function of the integrals can be efficiently approximated as

$$k_{\text{far}}(\vec{\boldsymbol{x}}, \vec{\boldsymbol{z}}_l) = g(\vec{\boldsymbol{x}})\, h(\vec{\boldsymbol{z}}_l) \tag{147}$$

and, thus,

$$\boldsymbol{y}_{\text{far}} = \boldsymbol{V} \cdot \boldsymbol{B} \cdot \boldsymbol{W} \cdot \boldsymbol{p}_k \tag{148}$$

with $\boldsymbol{V}$, $\boldsymbol{B}$ and $\boldsymbol{W}$ being sparse matrices.

The particular approximation of the kernel function in the boundary element formulation is somewhat more complex. Most popular is the kernel approximation based on a multipole series by the so-called spherical harmonics, see for example Schneider

(2003). Another technique is based on Chebyshev interpolation, cf. Messner et al. (2012). More specific techniques are possible and a large number of papers has been published about these over the last two decades.

While the simple fast multipole technique is reducing the complexity from $O(N^2)$ to $O(N^{1.5})$, the multilevel fast multipole method is able to further reduce the complexity. Introduction of a cluster tree which is required for a multilevel fast multipole method allows to reduce memory requirements close to $O(N)$ and complexity to $O(N \log N)$. Complexity depends on frequency as well.

8.3 Example: Music Recording Studio

Three Round Robins in room acoustics had been launched by the Physikalisch–Technische Bundesanstalt (PTB) in Braunschweig (Germany) between 1994 and 2000, cf. Bork (2000, 2005a, b). The third phase of the Round Robin considered the music recording studio of the PTB which accounts for the computational example in this section. The results have been published earlier in Marburg et al. (2003). The boundary element results were yielded using a multilevel fast multipole method which had been implemented into the computer code Akusta. These results are compared with experimental data and with simulation results stemming from the room acoustics computer code Caesar, cf. Vorländer (1989). It has been the objective of this example to show that results of a BEM simulation can be used to determine room acoustic measures.

The geometry model is based on the data given for the third phase of this third Round Robin. These data provided absorption coefficients α for all surfaces. A real boundary admittance was evaluated as

$$|\tilde{Y}| = \frac{1 - \sqrt{1-\alpha}}{1 + \sqrt{1-\alpha}}. \tag{149}$$

Further, it was assumed that the (closed) curtains are fixed at the windows. Obviously, this assumption is not realistic since the curtains are hanging some centimeters inside the room. However, the boundary element model requires this.

The frequency range for the octave bands 125, 250, 500 Hz, and 1 kHz were analyzed. The octave band of 1kHz required an analysis up to 1450 Hz. The studio geometry was essentially looking like a rectangular room of $8 \times 9\,\mathrm{m}^2$ and a height of 5 m. The longest distance in the model is 13.8 m. Hence, the maximum normalized wavenumber is found for $k_{\max} l_{\max} \approx 367$.

Three different boundary element models of constant elements were prepared for three different frequency ranges:

1. $h_{\max} \leq 20$cm: 9169 elements, $80 \ldots 400$ Hz, $kh = 0.29 \ldots 1.46$
2. $h_{\max} \leq 10$cm: 35553 elements, $360 \ldots 750$ Hz, $kh = 0.66 \ldots 1.37$
3. $h_{\max} \leq 6$cm: 98424 elements, $700 \ldots 1450$ Hz, $kh = 0.77 \ldots 1.59$

For the highest frequency of 1450 Hz, this accounts for just 4 boundary elements per wavelength. The computation encompassed the evaluation of six transfer functions, i.e. the transfer functions between two monopole sources and three receivers. These transfer functions are related to the sound pressure of the source in free space in 10 m distance.

Figure 28 shows an arbitrarily chosen transfer function for the entire frequency range over four octave bands, i.e. from 80 to 1450 Hz in steps of 0.5 Hz. In overlapping regions between 360 and 400 Hz as well as between 700 and 750 Hz, the results show good agreement. However, it is obvious that hardly anything else can be read out of this curve. Therefore, typical room acoustic quantities were evaluated. It is common to evaluate these quantities for octave bands only. Strength G and reverberation time T_{30} are displayed for two octave bands in Fig. 29. It becomes clear that the BEM results capture the specific differences between the six transfer functions rather well

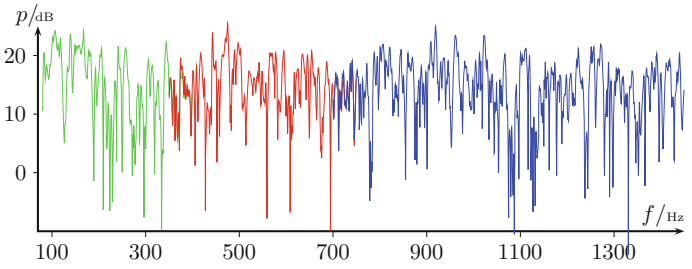


Fig. 28 Transfer function between Source 1 (S1) and Receiver 3 (R3) related to free-field source (10 m distance from receiver)

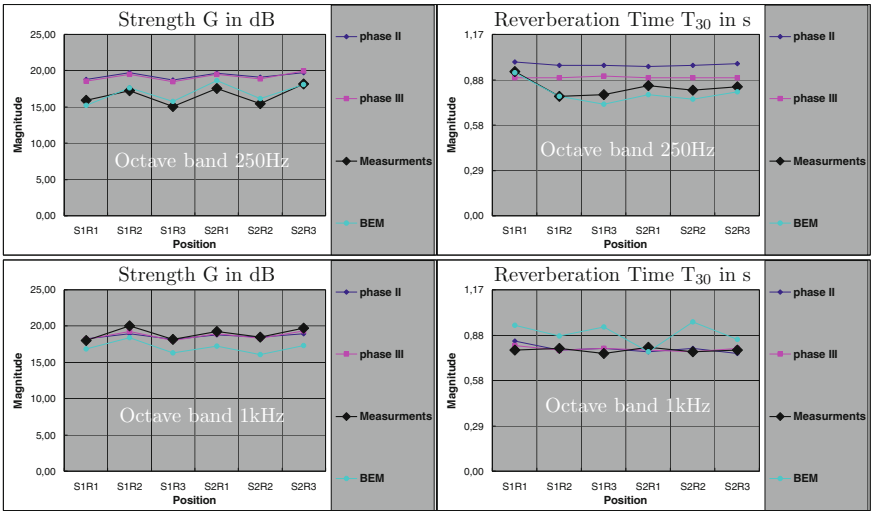


Fig. 29 Strength and reverberation time for two different octave bands and six transfer paths

in the lower frequency range while the room acoustics code is able to predict the shown quantities much better in the higher frequency range. Furthermore, the room acoustics code provides results much faster than the BEM code which solves the complicated boundary value problem over a wide frequency range with a very fine frequency resolution which is finally lost, when the octave band results are determined by integration over these results. Therefore, the author is convinced that evaluating the high frequency range with BEM is not reasonable since other methods do this better and are much faster. The availability of fast BEM techniques does not necessarily guarantee that its use is sensible. However, there may be cases for which large scale boundary element models are useful and practical. For the music recording studio, this was not the case.

9 Fluid-Structure Interaction

9.1 Coupled Systems of Equations

In the previous sections on the boundary element method, only wave propagation in fluids was considered. This section deals with the treatment of interaction between fluid and structure. Often, only a one-way coupling between structure and fluid is considered. In such a case, the structure is assumed to vibrate in vacuo and then, the structural vibrations are applied as boundary conditions, i.e. particle velocity, for the fluid. Somehow, the boundary admittance can represent the structural behavior as long as the structure is locally reacting which means that there should not be any sound propagation within the surrounding structure. A full fluid-structure interaction allows sound propagation in the structure and thus, a feedback from the fluid to the structure.

The interaction is achieved by two coupling conditions between fluid and structure. At first, it is assumed that the particle velocity at the interface in normal direction is continuous

$$\vec{\boldsymbol{v}}_f \cdot \vec{\boldsymbol{n}} - \vec{\boldsymbol{v}}_s \cdot \vec{\boldsymbol{n}} = 0 \qquad \text{or} \qquad v_f(\vec{\boldsymbol{x}}) - v_s(\vec{\boldsymbol{x}}) = 0. \tag{150}$$

A second condition is defined by balancing the momentum as

$$\vec{\vec{\boldsymbol{\sigma}}} \cdot \vec{\boldsymbol{n}} + p\,\vec{\boldsymbol{n}} = \vec{\boldsymbol{0}}, \tag{151}$$

where the surface traction in normal direction equals the sound pressure. The latter condition accounts for the one which is usually omitted for the one-way coupling. It is relevant for cases of light structures and/or heavy fluid. However, in cases for which damping by sound radiation is relevant, e.g. sound radiation from carillon bells (Roozen-Kroon 1992), even for a heavy structure and a light fluid, the interaction may not be negligible since sound radiation might be the dominant source of structural damping.

Herein, description of the coupling formulation is started with the balance of momentum, cf. Eq. (151), which results in an excitation of the structure by the sound pressure. This additional excitation appears as an additional load term in the system of equations for the structure

$$\left(\boldsymbol{C} - \mathrm{i}\omega\boldsymbol{B} - \omega^2\boldsymbol{M}\right)\boldsymbol{u}^s = \boldsymbol{A}\boldsymbol{u}^s = \boldsymbol{f}^s + \boldsymbol{C}^{sf}\,\boldsymbol{p}^f, \tag{152}$$

which is usually generated by a finite element discretization. In this formulation $\boldsymbol{C}$, $\boldsymbol{B}$ and $\boldsymbol{M}$ are the static stiffness matrix, the damping matrix and the mass matrix, respectively. $\boldsymbol{A}$ is known as the dynamic stiffness matrix of the structure, $\boldsymbol{u}^s$ and $\boldsymbol{f}^s$ contain the nodal data of structural displacements and the discrete forces acting on the structure both on the structural mesh, respectively, whereas $\boldsymbol{p}^f$ contains the nodal sound pressure values on the fluid mesh. $\boldsymbol{C}^{sf}$ is the coupling matrix between structure and fluid as

$$C^{sf}_{kj\,m} = \int_\Gamma \phi^s_k n^s_m \phi^f_j \, d\Gamma \tag{153}$$

with ϕ^s_k being the test function k for the structural displacement, ϕ^f_j as the interpolation function j for the sound pressure and n^s_m containing the components of the normal vector on the structural mesh since the fluid sound pressure results only in a normal traction of the structure.

For the fluid equation, we start from Eq. (66) and assume zero admittance because the admittance represents the fluid structure interaction as will be seen later. Equation (66) is rewritten as

$$\boldsymbol{H}\,\boldsymbol{p}^f - \boldsymbol{G}\boldsymbol{v}^f_s = \boldsymbol{f}^f \tag{154}$$

with $\boldsymbol{p}^f$ and $\boldsymbol{v}^f_s$ defined on the fluid mesh.

The straightforward approach couples the degrees of freedom directly at the nodes of coincident structural and fluid meshes as

$$\vec{\boldsymbol{u}}^s \cdot \vec{\boldsymbol{n}}^s - \mathrm{i}\omega v^f_s = 0. \tag{155}$$

In literature, this approach is often referred to as point-to-point coupling. Although simple and convenient, this approach is not recommended at all. It is not very accurate and requires coincident grids. The so-called mortar methods solve these problems, see for example Bernardi et al. (1994), Flemisch et al. (2006). For this approach, Eq. (155) is tested with the test functions of the fluid ϕ^f_l. Applying interpolation for $\vec{\boldsymbol{u}}^s$ and v^f_s yields

$$\int_\Gamma \phi^f_l \left(\sum_{j=1}^{N_s} \phi^s_j \vec{\boldsymbol{n}}^s_j \vec{\boldsymbol{u}}^s_j \right) d\Gamma - \mathrm{i}\omega \int_\Gamma \phi^f_l \left(\sum_{k=1}^{N_f} \phi^f_k v^f_{s_k} \right) d\Gamma = 0. \tag{156}$$

This is easily put into matrix form as

$$\boldsymbol{C}^{sf\,T}\boldsymbol{u}^s - \mathrm{i}\omega\boldsymbol{\Theta}\boldsymbol{v}_s^f = 0, \tag{157}$$

where $\boldsymbol{C}^{sf}$ is the same matrix as given in Eq. (153) and $\boldsymbol{\Theta}$ is again the boundary mass matrix of the fluid mesh as given in Eq. (34). The resulting system of equations now reads as

$$\begin{bmatrix} \boldsymbol{A} & -\boldsymbol{C}^{sf} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{H} & -\boldsymbol{G} \\ \boldsymbol{C}^{sf\,T} & \boldsymbol{0} & -\mathrm{i}\omega\boldsymbol{\Theta} \end{bmatrix} \begin{bmatrix} \boldsymbol{u}^s \\ \boldsymbol{p}^f \\ \boldsymbol{v}_s^f \end{bmatrix} = \begin{bmatrix} \boldsymbol{f}^s \\ \boldsymbol{f}^f \\ \boldsymbol{0} \end{bmatrix}. \tag{158}$$

This equation is easily simplified by isolation and elimination of $\boldsymbol{v}_s^f$

$$\boldsymbol{v}_s^f = -\frac{\mathrm{i}}{\omega}\boldsymbol{\Theta}^{-1}\boldsymbol{C}^{sf\,T}\boldsymbol{u}^s = \boldsymbol{C}^{fs}\boldsymbol{u}^s. \tag{159}$$

$\boldsymbol{C}^{fs}$ is representing the coupling matrix between fluid and structure. Thus, the system of equations for the interaction between structure and fluid is written as

$$\begin{bmatrix} \boldsymbol{A} & -\boldsymbol{C}^{sf} \\ -\boldsymbol{G}\boldsymbol{C}^{fs} & \boldsymbol{H} \end{bmatrix} \begin{bmatrix} \boldsymbol{u}^s \\ \boldsymbol{p}^f \end{bmatrix} = \begin{bmatrix} \boldsymbol{f}^s \\ \boldsymbol{f}^f \end{bmatrix}. \tag{160}$$

In this equation, the first row represents the structural equation whereas the second row can be understood as the fluid equation. As mentioned above, it is possible to couple very different meshes of different polynomial degree and size. A technique to couple even non-planar interfaces on different meshes was described in Peters et al. (2012).

In a further step, the structural displacement $\boldsymbol{u}^s$ in the first row of Eq. (160) is first isolated as

$$\boldsymbol{u}^s = \boldsymbol{A}^{-1}\left(\boldsymbol{f}^s + \boldsymbol{C}^{sf}\boldsymbol{p}^f\right) \tag{161}$$

and then replaced in the fluid equation, which can be rewritten as

$$\left(\boldsymbol{H} - \boldsymbol{G}\underbrace{\boldsymbol{C}^{fs}\boldsymbol{A}^{-1}\boldsymbol{C}^{sf}}_{\boldsymbol{Y}_c}\right)\boldsymbol{p}^f = \boldsymbol{f}^f + \boldsymbol{G}\boldsymbol{C}^{fs}\boldsymbol{A}^{-1}\boldsymbol{f}^s. \tag{162}$$

Hence, the system of equations has the same structure as in Eqs. (28), (35), (66) and (73)

$$\left(\boldsymbol{H} - \boldsymbol{G}\boldsymbol{Y}_c\right)\boldsymbol{p}^f = \boldsymbol{f}^{fs} \tag{163}$$

which means that $\boldsymbol{Y}_c$ can be understood as a coupling admittance matrix, cf. Fritze et al. (2005), Marburg and Anderssohn (2011).

9.2 Example: Long Duct with Multibody System

For the demonstrating example, we return to the 1d duct problem with an attached three degrees of freedom multibody system. This example is well suited to explain how the boundary admittance represents the fluid structure interaction. For numerical examples including a comparison with analytic solutions, the reader is referred to, among others, Peters et al. (2012). In further manuscripts of these authors, the algebraic eigenvalue problem for coupled models in exterior acoustics is set up, solved and compared with analytic solutions, cf. Peters et al. (2013, 2014).

The configuration of the duct problem with an attached three degrees of freedom multibody system is shown in Fig. 30. It will be distinguished between the fluid problem in the duct and the multibody system around the duct. The structural behaviour is governed by a set of equations of motion which will be derived below. However, we will start with recalling the fluid matrices. This example has been presented in Marburg and Anderssohn (2011). The results herein are similar. They differ in their form.

Fluid Matrices For the fluid, only the matrices $\boldsymbol{G}$ and $\boldsymbol{H}$ are relevant. The matrices are given in Eq. (51). The fluid part of the system of equations is the same as in Eq. (154) but with vanishing right hand side and unknown boundary data (p_0, p_l, v_0, v_l).

Equations of Motion for Multibody System For formulation of the equations of motion of the multibody system, it is useful to analyse the problem first. The structural part of the system in Fig. 30 consists of three bodies with one degree of freedom each. The three bodies with masses of m_0, m_l and m_b are allowed to move horizontally

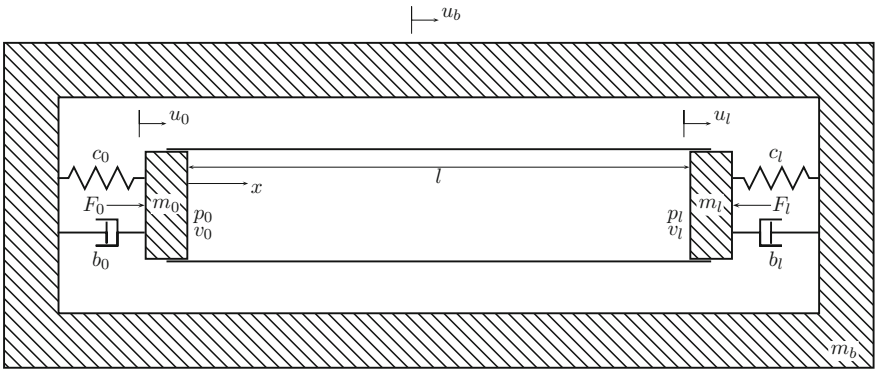


Fig. 30 Configuration of duct problem with discrete three degrees of freedom multibody system attached

and their movements are described by the translational coordinates u_0, u_l and u_b, respectively. The outer box connects the two end caps of the duct. The three equations of motion in frequency domain are written as

$$\begin{gathered}\left(-\omega^2 m_0 - \mathrm{i}\omega b_0 + c_0\right) u_0 + \left(\mathrm{i}\omega b_0 - c_0\right) u_b = \delta_0 u_0 - \gamma_0 u_b = F_0 - A p_0, \\ \left(-\omega^2 m_l - \mathrm{i}\omega b_l + c_l\right) u_l + \left(\mathrm{i}\omega b_l - c_l\right) u_b = \delta_l u_l - \gamma_l u_b = -F_l + A p_l, \\ \left[-\omega^2 m_b - \mathrm{i}\omega\left(b_0 + b_l\right) + \left(c_0 + c_l\right)\right] u_b + \left[\mathrm{i}\omega b_0 - c_0\right] u_0 + \left[\mathrm{i}\omega b_l - c_l\right] u_l = 0\end{gathered} \tag{164}$$

for the left and the right end cap and for the outer box, respectively. The latter is rewritten in an abridged form as

$$\delta_b u_b - \gamma_0 u_0 - \gamma_l u_l = 0\,. \tag{165}$$

In Eq. (164), b_k and c_k are the damping coefficient and the spring's stiffness, respectively. F_k represents the excitation force and index k can be either 0 or l depending on the end of the duct which is considered. A is the cross sectional area of the duct. It is needed to determine the acoustic excitation force acting on the end caps. In a next step, u_b is isolated as

$$u_b = \frac{\gamma_0 u_0 + \gamma_l u_l}{\delta_b} \tag{166}$$

and replaced. Then, the first two equations of motion in Eq. (164) change into

$$\begin{gathered}\left(\delta_0 - \frac{\gamma_0^2}{\delta_b}\right) u_0 - \frac{\gamma_0 \gamma_l}{\delta_b} u_l = F_0 - A p_0 \quad \text{and} \\ -\frac{\gamma_0 \gamma_l}{\delta_b} u_0 + \left(\delta_l - \frac{\gamma_l^2}{\delta_b}\right) u_l = -F_l + A p_l\,.\end{gathered} \tag{167}$$

With these formulations, we fill the dynamic stiffness matrix $\boldsymbol{A}$ and adjust Eq. (152) in order to produce

$$\begin{aligned}\boldsymbol{A}\boldsymbol{u} &= \begin{bmatrix} \delta_0 - \frac{\gamma_0^2}{\delta_b} & -\frac{\gamma_0 \gamma_l}{\delta_b} \\ -\frac{\gamma_0 \gamma_l}{\delta_b} & \delta_l - \frac{\gamma_l^2}{\delta_b} \end{bmatrix} \begin{bmatrix} u_0 \\ u_l \end{bmatrix} = \\ &= \begin{bmatrix} F_0 \\ -F_l \end{bmatrix} + \begin{bmatrix} -A & 0 \\ 0 & A \end{bmatrix} \begin{bmatrix} p_0 \\ p_l \end{bmatrix} = \boldsymbol{f} + \boldsymbol{C}_{sf}\,\boldsymbol{p}\,.\end{aligned} \tag{168}$$

Now knowing the structural part, we are able to arrange the coupled system of equations into the form of Eq. (160). However, we still need the coupling matrix $\boldsymbol{C}_{fs}$. This is easily formulated since the particle velocity is related to the structural displacement as

$$\boldsymbol{v}_f = \begin{bmatrix} v_0 \\ v_l \end{bmatrix} = \begin{bmatrix} \mathrm{i}\omega & 0 \\ 0 & -\mathrm{i}\omega \end{bmatrix} \begin{bmatrix} u_0 \\ u_l \end{bmatrix} = \boldsymbol{C}_{fs}\boldsymbol{u}\,. \tag{169}$$

Hence, we can write

$$\begin{bmatrix} h_{11} & h_{12} & \mathrm{i}\omega g_{11} & -\mathrm{i}\omega g_{12} \\ h_{21} & h_{22} & \mathrm{i}\omega g_{21} & -\mathrm{i}\omega g_{22} \\ A & 0 & \delta_0 - \frac{\gamma_0^2}{\delta_b} & -\frac{\gamma_0\gamma_l}{\delta_b} \\ 0 & -A & -\frac{\gamma_0\gamma_l}{\delta_b} & \delta_l - \frac{\gamma_l^2}{\delta_b} \end{bmatrix} \begin{bmatrix} p_0 \\ p_l \\ u_0 \\ u_l \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ F_0 \\ -F_l \end{bmatrix}. \tag{170}$$

In order to replace the structural degrees of freedom by applying the Schur complement, inversion of matrix $\boldsymbol{A}$ is required. This yields

$$\boldsymbol{A}^{-1} = \begin{bmatrix} \tilde{a}_{11} & \tilde{a}_{12} \\ \tilde{a}_{21} & \tilde{a}_{22} \end{bmatrix} = \frac{1}{N} \begin{bmatrix} \delta_l - \frac{\gamma_l^2}{\delta_b} & \frac{\gamma_0\gamma_l}{\delta_b} \\ \frac{\gamma_0\gamma_l}{\delta_b} & \delta_0 - \frac{\gamma_0^2}{\delta_b} \end{bmatrix} \tag{171}$$

with

$$N = \delta_0\delta_l - \frac{\delta_0\gamma_l^2}{\delta_b} - \frac{\delta_l\gamma_0^2}{\delta_b}. \tag{172}$$

In analogy to Eq. (162), application of the Schur complement yields

$$\begin{aligned} (\boldsymbol{H} - \boldsymbol{G}\boldsymbol{Y}_c)\,\boldsymbol{p} &= \left\{ \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} - \mathrm{i}\omega A \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \begin{bmatrix} -\tilde{a}_{11} & \tilde{a}_{12} \\ \tilde{a}_{21} & -\tilde{a}_{22} \end{bmatrix} \right\} \begin{bmatrix} p_0 \\ p_l \end{bmatrix} = \\ &= \mathrm{i}\omega \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \begin{bmatrix} \tilde{a}_{11}F_0 - \tilde{a}_{12}F_l \\ -\tilde{a}_{21}F_0 + \tilde{a}_{22}F_l \end{bmatrix} = \boldsymbol{G}\boldsymbol{C}^{fs}\boldsymbol{A}^{-1}\boldsymbol{f}^s. \end{aligned} \tag{173}$$

Thus, we extract the fully populated admittance matrix $\boldsymbol{Y}_c$ as

$$\boldsymbol{Y}_c = \mathrm{i}\omega A \begin{bmatrix} -\tilde{a}_{11} & \tilde{a}_{12} \\ \tilde{a}_{21} & -\tilde{a}_{22} \end{bmatrix} = \frac{-\mathrm{i}\omega A}{N} \begin{bmatrix} \delta_l - \frac{\gamma_l^2}{\delta_b} & -\frac{\gamma_0\gamma_l}{\delta_b} \\ -\frac{\gamma_0\gamma_l}{\delta_b} & \delta_0 - \frac{\gamma_0^2}{\delta_b} \end{bmatrix} \tag{174}$$

with the constants

$$\begin{aligned} \gamma_0 &= -\mathrm{i}\omega b_0 + c_0 \\ \gamma_l &= -\mathrm{i}\omega b_l + c_l \\ \delta_0 &= -\omega^2 m_0 - \mathrm{i}\omega b_0 + c_0 \\ \delta_l &= -\omega^2 m_l - \mathrm{i}\omega b_l + c_l \\ \delta_b &= -\omega^2 m_b - \mathrm{i}\omega\,(b_0 + b_l) + (c_0 + c_l) \\ N &= \delta_0\delta_l - \frac{\delta_0\gamma_l^2}{\delta_b} - \frac{\delta_l\gamma_0^2}{\delta_b}. \end{aligned} \tag{175}$$

It is not straightforward to check correctness of these results for the admittance matrix. At least, the solution passes the plausibility check for the case that the box mass m_b becomes large, i.e. $m_b \to \infty$. Then, the constants $\delta_b \to \infty$, hence $1/\delta_b = 0$. N reduces to $N = \delta_0 \delta_l$. This means that for $m_b \to \infty$, $\boldsymbol{Y}_c$ becomes a diagonal matrix with admittance values for the single degree of freedom model. In that case, and in particular for the left end cap, the admittance is

$$Y_0 = -\frac{\mathrm{i}\omega A}{\delta_0} = \frac{-\mathrm{i}\omega A}{-\omega^2 m_0 - \mathrm{i}\omega b_0 + c_0}. \tag{176}$$

A diagonal admittance matrix represents a locally reacting boundary, i.e. there is no sound (and vibration) propagation in the boundary. Furthermore, this means that a purely real admittance value represents a dashpot whereas a purely imaginary admittance assumes that there is just an undamped single degree of freedom system which consists of a spring and a rigid body with a certain mass. This is a reasonable result.

10 Symmetric (Halfspace) and Periodic Problems

10.1 Formulation

Many technical problems in acoustics involve symmetries and periodicity. A typical problem is the radiation from baffled plates and shells. A baffled structure is a structure which radiates into a halfspace, whereas the ground plane of the halfspace acts as a symmetry plane, i.e. it assumed to be acoustically rigid. Herein, it is assumed that the plane $z = 0$ is the symmetry plane. The formulation can easily be extended to other symmetry planes and to multiple symmetry. While the symmetric case assumes $p(z) = p(-z)$, the asymmetric case $p(z) = -p(-z)$ is also relevant and rather easy to consider, see for example Wu and Seybert (1991). Periodic boundary element techniques have become popular in recent years in particular for the analysis of noise barriers and sonic crystals, see for example Lam (1999), Jean and Defrance (2015), Fard et al. (2015, 2017), Karimi et al. (2016, 2017), Ziegelwanger et al. (2017).

For the symmetric case, it is assumed that geometry, boundary conditions, incident sound field (if applicable) and the resulting sound field are symmetric, cf. configurations in Fig. 31. The asymmetric case assumes that the geometry is symmetric and the boundary conditions are only symmetric with respect to the x and the y components but asymmetric for the z component. Then, the sound pressure will be zero in the symmetry plane and, as mentioned above, $p(z) = p(-z)$. The periodic case is sketched in Fig. 32. There, the geometry, the boundary conditions, the incident sound field (if applicable) and the resulting sound field are periodic.

Boundary element techniques allow for a straightforward consideration of symmetry and periodicity and, in contrast to the finite element method, do not require adjustment of boundary conditions at interfaces. In order to derive a suitable formulation, it is useful to recall Eqs. (23)–(26) which represent the Kirchhoff–Helmholtz integral equation discretized by collocation. Simplification of these equations is yielded by assuming Neumann boundary conditions, i.e. $Y = 0$

$$c(\vec{z}_l)\, p_l + \int_\Gamma \frac{\partial G(\vec{x}, \vec{z}_l)}{\partial n(\vec{x})}\, p(\vec{x}) d\Gamma(\vec{x}) = sk \int_\Gamma G(\vec{x}, \vec{z}_l)\, v_s(\vec{x}) d\Gamma(\vec{x}). \tag{177}$$

In this equation, the collocation point is assumed to be either above the symmetry plane in Fig. 31 or within the k-th segment as shown in Fig. 32. Since the sound pressure is either assumed to be symmetric or periodic, the integrals can be split up as

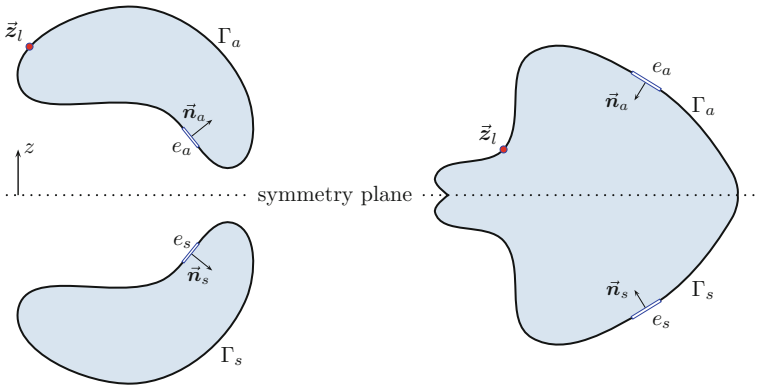


Fig. 31 Configuration of halfspace (symmetric) problems

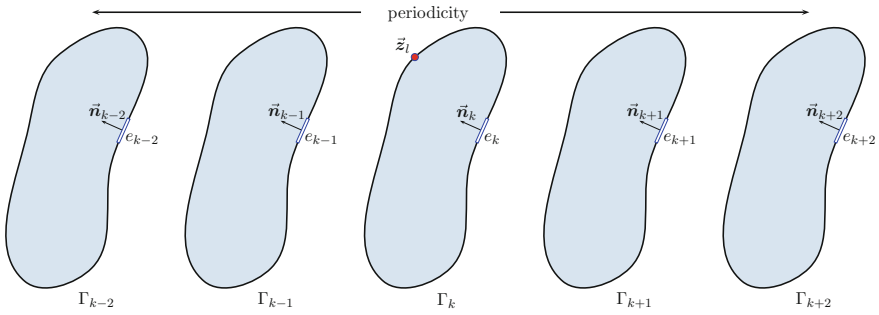


Fig. 32 Configuration of periodic problems

$$c(\vec{z}_l)\, p_l + \sum_{i=1}^{N_s} \int_{\Gamma_i} \frac{\partial G(\vec{x}, \vec{z}_l)}{\partial n(\vec{x})} p(\vec{x}) d\Gamma_i(\vec{x}) = sk \sum_{i=1}^{N_s} \int_{\Gamma_i} G(\vec{x}, \vec{z}_l)\, v_s(\vec{x}) d\Gamma_i(\vec{x}) \tag{178}$$

with N_s being the number of segments considered. For the symmetric halfplane problems in Fig. 31, $N_s = 2$. For the periodic problem sketched in Fig. 32, $N_s \to \infty$ which is not very practical. Therefore, it was suggested in Fard et al. (2015) to use a finite N_s which makes the method only a quasi-periodic boundary element method. Several hundreds of segments may be required though. The introduction of interpolation functions and assembly into matrices leads to the system of equations as

$$\left[\sum_{i=1}^{N_s} \boldsymbol{H}_i\right] \boldsymbol{p} = \left[\sum_{i=1}^{N_s} \boldsymbol{G}_i\right] \boldsymbol{v}_s \quad \longrightarrow \quad \boldsymbol{H}\boldsymbol{p} = \boldsymbol{G}\boldsymbol{v}_s. \tag{179}$$

In case of the periodic model, there are no adjustments required for the different segments. In case of symmetry, however, there are a few adjustments necessary. When using the straightforward approach as shown here, the user needs to turn the normal vectors into the opposite z-direction. If the asymmetric case is considered, the signs in front of the surface integrals over Γ_s must be adjusted suitably.

Often, it is more efficient to use a modified Green's function to consider the different segments and to avoid integrating over all segments one after the other. Instead of keeping the collocation point just in one position and integrating over all segments one after the other, it is more efficient to use the principle of reciprocity. In practice, this means that a collocation point is assumed to be located on each segment at the same position and only one integration over the primary segment is required. Such an approach is common practice for symmetric and asymmetric problems as shown in Wu and Seybert (1991).

Other authors took advantage of the matrix structure which is observed for periodic models, cf. Karimi et al. (2016, 2017). These matrices show a clear Toeplitz structure which can be used to efficiently reduce the memory requirements and even enable the user to solve for a non-periodic sound pressure on a periodic geometry.

The quasi-periodic technique presented above performs nicely together with a fast boundary element technique. This has been demonstrated in Ziegelwanger et al. (2017). As a practical application, it is well suited to accommodate different geometric structures along the length of the barrier. Typical elements are differently tuned Helmholtz resonators as demonstrated in Fard et al. (2017).

10.2 Example: Monopole Source Close to Tire and Road

The example which is briefly discussed here has already been presented in a former paper of the author, cf. Marburg et al. (2002). For this, we consider the outer surface of an entire sedan on a rigid plane, i.e. the half-space problem is solved. In order to

avoid the effect of irregular frequencies, the Burton and Miller method is applied and the vehicle surface is assumed to be rigid, i.e. $Y = 0$. A monopole excitation very close to and in front of the left rear tire is applied to simulate noise radiated from the tire. The boundary element mesh consists of 25808 constant elements which have an edge length of maximum 6 cm. Analysis is limited to a frequency range up to 1100 Hz in air. The problem has been solved in the steps of 1.0 Hz to capture all resonance peaks sufficiently accurate. The mesh is refined in the wheel house regions and a discontinuous topology has been allowed to keep the degree of freedom small. Since the symmetry plane is not meshed in the boundary element method, the regions of the contact patches appear as holes in Fig. 33.

Figure 34 presents noise transfer function as the difference between the calculated sound pressure level and a free monopole above the symmetry plane. This noise transfer function and, also, most other noise transfer functions which had been investigated

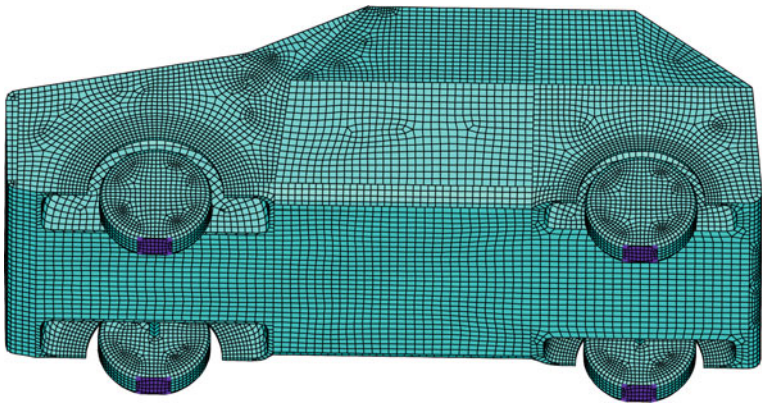


Fig. 33 Boundary element model of a car for the tire noise problem in halfspace. Excitation by monopole in front of and close to *left* rear tire. *Holes* in the tires show the contact patches

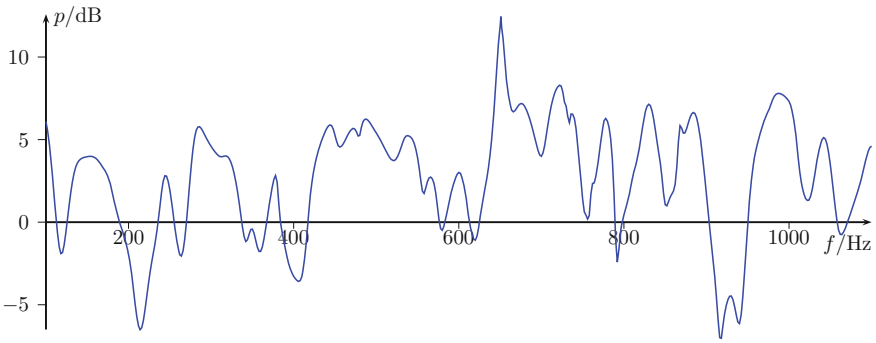


Fig. 34 Noise transfer function for point 1.5 m *left* of *left* rear tire, sound pressure due to monopole source related to free monopole source above the symmetry plane

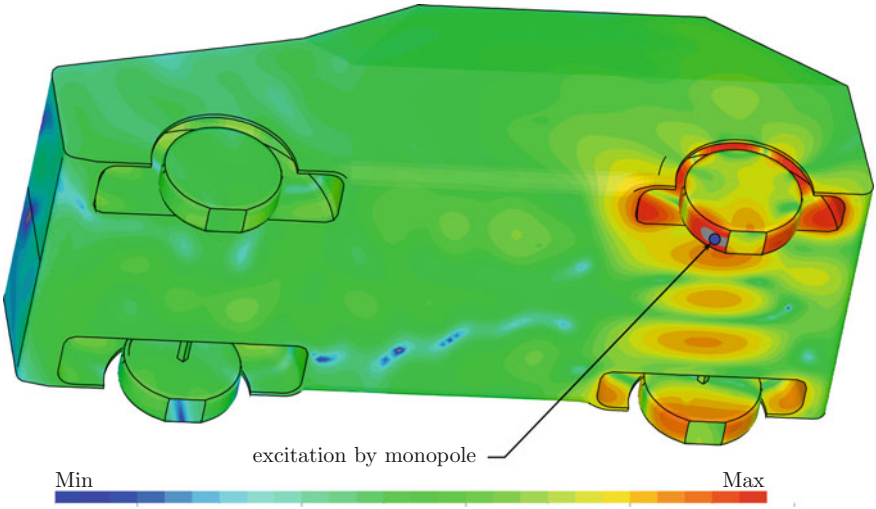


Fig. 35 Surface sound pressure level distribution due to monopole excitation at *left* rear tire for a frequency of 651 Hz

show numerous peaks which clearly correspond to resonances. These resonances are identified as standing wave phenomena in the wheel houses and between the tires. They occur due to weakly damped modes of open resonators, cf. Marburg et al. (2006), Marburg (2006), which are utilized as musical instruments, cf. Fuß et al. (2011), Moheit and Marburg (2017). However, the eigenvalue problem has not been solved in this case.

The most distinguished operational mode shape and corresponding resonance peak is observed for the frequency of 651 Hz. The surface sound pressure distribution for 651 Hz is shown in Fig. 35. The vibration mode exhibits five maxima and four nodes along the semi-circle of the tire housing. Additionally, another mode of three maxima between both rear tires is identified. The same mode as in the left tire housing is found in the right one but with lower amplitudes.

11 Panel Contribution Analysis

So far, the analysis in this chapter was only focusing on the solution of boundary value problems. In industrial applications, however, the engineers encounter the problem to identify the surface regions which contribute the most to a specified objective function, e.g. the sound pressure (level) at certain discrete points or the radiated sound power, cf. Marburg (2002a), Marburg et al. (2016). While it is possible to do panel contribution analysis using other methods too, the boundary element formulation

allows a very easy derivation of this. Nowadays, most commercial finite and boundary element codes offer a feature which is called panel contribution analysis.

Panel contribution analysis is also a feature of software for experimental equipment. Two recent papers combine a method of inverse acoustics with panel contribution analysis, cf. Wu and Natarajan (2013), Wu et al. (2015).

11.1 Surface Contributions with Respect to the Sound Pressure

A first and simple concept was suggested by Ishiyama et al. (1988) for surface contributions to the sound pressure at a single point. For this, the derivation is started with recalling Eq. (81) as

$$p(\vec{\boldsymbol{y}}) \;=\; \boldsymbol{g}^T(\vec{\boldsymbol{y}})\boldsymbol{v}_s - \left[\boldsymbol{h}^T(\vec{\boldsymbol{y}}) - \boldsymbol{g}^T(\vec{\boldsymbol{y}})\boldsymbol{Y}\right]\boldsymbol{p}. \tag{180}$$

Substituting the nodal sound pressure values $\boldsymbol{p}$ by application of Eq. (28) and $\boldsymbol{D} = \boldsymbol{GY}$ yields

$$p(\vec{\boldsymbol{y}}) = \underbrace{\left\{\boldsymbol{g}^T(\vec{\boldsymbol{y}}) - \left[\boldsymbol{h}^T(\vec{\boldsymbol{y}}) - \boldsymbol{g}^T(\vec{\boldsymbol{y}})\boldsymbol{Y}\right](\boldsymbol{H} \;-\; \boldsymbol{GY})^{-1}\,\boldsymbol{G}\right\}}_{\boldsymbol{b}^T(\vec{\boldsymbol{y}})}\boldsymbol{v}_s. \tag{181}$$

Column matrix $\boldsymbol{b}$ can be understood as the sensitivity of the sound pressure at a field point with respect to the surface particle velocity, cf. Coyette et al. (1993), Adey et al. (1995), Dong et al. (2004). In literature, this column matrix is also referred to as acoustic transfer vector, cf. Cremers et al. (2000), or as discrete acoustic influence coefficients, cf. Marburg et al. (1997), Marburg (2002a). Acoustic influence coefficients are determined as

$$\boldsymbol{b}(\vec{\boldsymbol{y}}) = \boldsymbol{g}(\vec{\boldsymbol{y}}) - \boldsymbol{G}^T\,(\boldsymbol{H} \;-\; \boldsymbol{GY})^{-T}\left[\boldsymbol{h}(\vec{\boldsymbol{y}}) - \boldsymbol{Y}^T\boldsymbol{g}(\vec{\boldsymbol{y}})\right], \tag{182}$$

in which $()^{-T}$ represents the inverse of a transpose matrix. Therefore, the formulation of influence coefficients can be understood as an adjoint operator approach for sensitivity analysis.

The scalar product between the column vectors of these influence coefficients and the particle velocities can be rewritten as

$$p(\vec{\boldsymbol{y}}) = \boldsymbol{b}^T(\vec{\boldsymbol{y}})\boldsymbol{v}_s = \sum_{k=1}^{N_n} b_k(\vec{\boldsymbol{y}})v_{sk} = \sum_{k=1}^{N_n}\eta_k \tag{183}$$

where η_k is the contribution of node k to the sound pressure at $\vec{\boldsymbol{y}}$. Summation of the these nodal contributions over a certain surface panel provides the user with the

information about the contribution of a particular panel to the overall sound pressure at that point.

Although looking quite elegant at first glance, this approach has some disadvantages. Since both, the influence coefficients and the particle velocity are arbitrarily complex and thus, might be erasing each other, a single panel contribution can be larger than the sound pressure at this particular point. It is easily possible to imagine the situation that two panels vibrate with a particular phase angle such that their contributions are of the same size with just 180 degrees phase difference. Then, these are large panel contributions with a vanishing sound pressure. Therefore, it is useful to handle these panel contributions very careful.

The nodal contributions η_k can be useful to conclude on surface activities relevant for the sound pressure at these particular points. This may be done by looking at a contour plot which visualizes the nodal contribution on the surface. However, these nodal contributions η_k as given in Eq. (183) are discrete values depending on the local discretization. Although difficult to understand at first, it may become clearer when looking at the particular formulation. Assume the scalar products of Eq. (183) for a coarse and a finer mesh with a constant surface particle velocity. The sound pressure at the field point must be independent of the mesh. However, the fine mesh consists of more nodes and elements than the coarse mesh. Consequently, the influence coefficients of the fine mesh must be smaller than those of the coarse mesh.

It is possible to visualise these discrete influence coefficients in contour plots by introducing a new magnitude which will be called continuous influence coefficients. These continuous influence coefficients β are then multiplied by (continuous) particle velocity to conclude on continuous surface contributions. The new continuous quantity is found by writing the sound pressure as

$$p(\vec{\boldsymbol{y}}) = \boldsymbol{b}^T(\vec{\boldsymbol{y}})\boldsymbol{v}_s = \int_\Gamma \beta(\vec{\boldsymbol{y}}) v_s d\Gamma. \tag{184}$$

The continuous boundary data on the right hand side, β and v_s, are approximated by the interpolation functions φ_j in the same way as discussed in the context of Eqs. (20) and (21). The nodal values of the continuous influence coefficients are assembled in $\boldsymbol{\beta}$. This results in

$$\begin{aligned} \boldsymbol{b}^T(\vec{\boldsymbol{y}})\,\boldsymbol{v}_s &= \int_\Gamma \left(\sum_{k=1}^{N} \varphi_k \beta_k(\vec{\boldsymbol{y}})\right)\left(\sum_{l=1}^{N} \varphi_l v_{sl}\right) d\Gamma \\ &= \sum_{k=1}^{N}\sum_{l=1}^{N} \beta_k(\vec{\boldsymbol{y}}) \underbrace{\left[\int_\Gamma \varphi_k \varphi_l d\Gamma\right]}_{\theta_{kl}} v_{sl} \end{aligned} \tag{185}$$

which can be written in matrix form as

$$\boldsymbol{b}^T(\vec{\boldsymbol{y}})\boldsymbol{v}_s = \boldsymbol{\beta}^T(\vec{\boldsymbol{y}})\boldsymbol{\Theta}\boldsymbol{v}_s. \tag{186}$$

Again, the sparse symmetric matrix $\boldsymbol{\Theta}$ is introduced. It is the same boundary mass matrix as introduced in Eq. (34). This matrix is used to switch between discrete and continuous quantities on the surface as

$$\boldsymbol{b}(\vec{\boldsymbol{y}}) = \boldsymbol{\Theta}\boldsymbol{\beta}(\vec{\boldsymbol{y}}) \tag{187}$$

which is a very similar application of the boundary mass matrix as for the coupling condition for fluid and structure, cf. Eq. (157).

Although providing a deeper insight into source identification, these surface contributions as formulated above have further limitations. It is not just that they can be arbitrarily complex and thus, extinguishing each other. It has hardly been discussed how to handle them over frequency ranges. A very simple (but also questionable) approach was presented by Marburg and Hardtke (2003) and even some example pictures have been shown therein. However, this problem, i.e. how to handle frequency ranges in this context, should receive further attention in the future.

A sedan cabin compartment model has been investigated. It is described in more detail in Marburg and Hardtke (2003). Originally, this model was created to investigate and optimize the structural model of the spare wheel well. It comprises cabin and trunk without a wall between them. This model includes a uniform boundary admittance which stems from measurements of the reverberation time, cf. Marburg and Hardtke (1999).

Figures 36 and 37 visualize influence coefficients with respect to the sound pressure at the driver's ear. This solution is equivalent to the surface contributions for a unit and uniform particle velocity applied to the entire surface of this sedan cabin compartment. Furthermore, it shows the sensitivity of the sound pressure at these particular frequencies (120 and 170 Hz) with respect to the surface particle velocity.

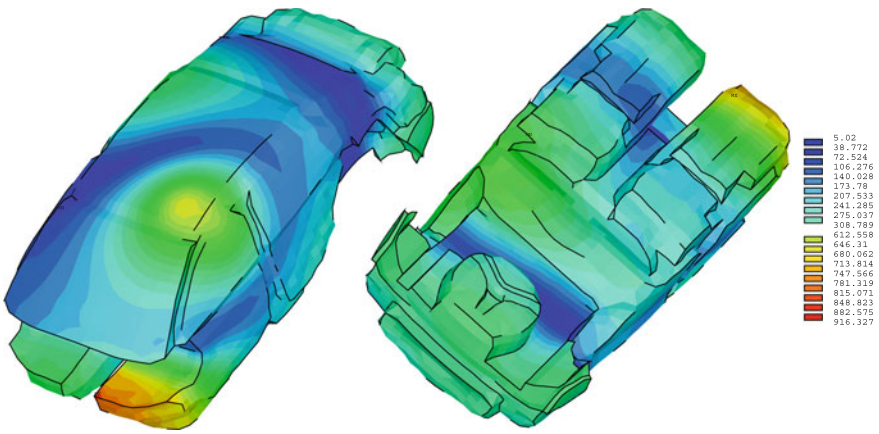


Fig. 36 Visualized influence coefficients which are equivalent to the surface contributions for a uniform particle velocity applied to the entire surface of this sedan cabin compartment (Frequency: 120 Hz)

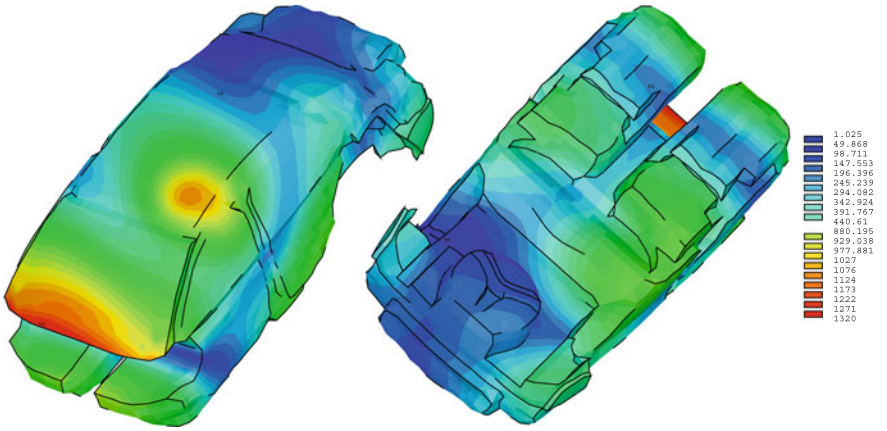


Fig. 37 Visualized influence coefficients which are equivalent to the surface contributions for a uniform particle velocity applied to the entire surface of this sedan cabin compartment (Frequency: 170 Hz)

It can be concluded that, for 120 Hz, the surface sensitivity is the highest close to the driver's head and in the region of the driver's feet. At the higher frequency of 170 Hz, the most sensitive region has moved to the windscreen and the upper dashboard. Even these two figures show that it is necessary to develop concepts to consider frequency ranges when discussing panel contribution analysis.

11.2 Contributions with Respect to Radiated Sound Power

While the sound pressure is quite popular as an objective function for cavities, sound radiation is usually evaluated based on the radiated sound power. The radiated sound power is defined as the (closed) surface integral over the acoustic intensity in normal direction. The closed surface can be an arbitrary enveloping surface either in the far or in the near field. In the context of boundary element formulations, it is most convenient to use the radiator's surface. Hence, we can write the radiated sound power P as

$$P = \frac{1}{2}\int_{\Gamma} \vec{\boldsymbol{I}} \cdot \vec{\boldsymbol{n}} d\Gamma = \frac{1}{2}\int_{\Gamma} \Re\left\{p v_f^*\right\} d\Gamma. \tag{188}$$

The straightforward approach for panel contribution analysis would be to integrate intensity over certain panels. A visualization of nodal intensity over the surface requires the use of the boundary mass matrix again and is easily displayed. In contrast to the surface contributions β in the previous subsection, intensity is real. However, it can be positive and negative. So, even for the acoustic short circuit, it looks as

if there is much activity and substantial surface contribution to the radiated sound power.

In contrast to the acoustic intensity, it is possible to formulate a non-negative type of intensity as shown by Marburg et al. (2013). In that paper, this quantity was called surface contribution but rather soon thereafter, Williams (2013) called it non-negative intensity.

The approach to this is described in what follows. We start with the discretized version of the radiated sound power as given in Eq. (87)

$$P = \frac{1}{2}\Re\left\{\boldsymbol{p}^T\boldsymbol{\Theta}\boldsymbol{v}_f^*\right\}. \tag{189}$$

Then, assuming $Y = 0$ for simplicity, the nodal surface sound pressure is replaced according to Eq. (28) which yields

$$P = \frac{1}{2}\Re\left\{\boldsymbol{v}_s^T\underbrace{\boldsymbol{G}^T\boldsymbol{H}^{-T}\boldsymbol{\Theta}}_{\tilde{\boldsymbol{Z}}^T}\boldsymbol{v}_s^*\right\} \tag{190}$$

where $\tilde{\boldsymbol{Z}}^T$ is known as the complex coupling matrix or, also, complex impedance matrix. Chen and Ginsberg (1995) have shown that this equation can be simplified into

$$P = \frac{1}{2}\boldsymbol{v}_s^T\Re\left\{\tilde{\boldsymbol{Z}}\right\}\boldsymbol{v}_s^* = \frac{1}{2}\boldsymbol{v}_s^T\boldsymbol{Z}\boldsymbol{v}_s^* \tag{191}$$

because of its symmetry properties. In practice, however, neither $\tilde{\boldsymbol{Z}}$ nor $\boldsymbol{Z}$ are actually symmetric. It is possible to symmetrize them since these matrices are converging to symmetric matrices. This was demonstrated by Peters and Kessissoglou (2012a).

Since $\boldsymbol{Z}$ is a real symmetric matrix, the square root of $\boldsymbol{Z}$ exists and can be determined using an eigenvalue decomposition. This eigenvalue decomposition results in the eigenvectors which are known as the radiation modes, cf. Marburg et al. (2013). With the knowledge of the square root of $\boldsymbol{Z}$ we can write

$$P = \frac{1}{2}\boldsymbol{v}_s^T\sqrt{\boldsymbol{Z}}\sqrt{\boldsymbol{Z}}^T\boldsymbol{v}_s^* \tag{192}$$

and then,

$$P = \frac{1}{2}\boldsymbol{\eta}^T\boldsymbol{\eta}^*. \tag{193}$$

Now, the radiated sound power is represented by the scalar product of a column matrix $\boldsymbol{\eta}$ with its complex conjugate. Writing it as a sum yields

$$P = \frac{1}{2}\sum_{k=1}^{N_n}\eta_k\eta_k^* = \frac{1}{2}\sum_{k=1}^{N_n}b_k \tag{194}$$

with all summands $b_k \geq 0$ and real numbers. Again, these b_k are discrete, i.e. mesh dependent, quantities. In a similar way as in the previous subsection, it is possible to conclude on a continuous quantity β which is well suited for visualization. This quantity β can be interpreted as an intensity which cannot be negative. Therefore, it was called non-negative intensity in Williams (2013).

These non-negative surface contributions (or non-negative intensities) are closely related to another quantity which has been subject of research for more than two decades now: the supersonic intensity, see for example Williams (1995), Magalhaes and Tenenbaum (2006), Fernandez-Grande et al. (2012). Both quantities have been compared to each other in a recent paper where it was shown that, when choosing the right parameters for the supersonic intensity, it results in the non-negative one and gives identical results. The approach shown here, however, is independent of such a parameter choice and therefore, most likely, more robust than the other approaches mentioned, cf. Liu et al. (2016a).

Another recent application of non-negative surface contributions presents a formulation to conclude on surface contributions to the scattered sound field including contributions to a certain far-field direction only, cf. Liu et al. (2016b).

The example in this subsection is an open resonator, which was presented in a former paper, cf. Peters et al. (2013a). This resonator consists of two square and parallel plates which are rigid and meshed by boundary elements. Both plates have a top surface area of 1.5 m^2 and are 0.915 m apart. The lower plate is 0.4 m thick and is fixed. The upper plate is 0.3 m thick and is flexibly mounted. It oscillates in vertical direction as a rigid piston with a particle velocity of 1 m/s.

Two configurations are considered. In the first one, these two plate pistons are just as described above. In the second configuration, a Helmholtz resonator is added to the fixed lower plate, cf. Fig. 38. Since Helmholtz resonators perform as tuned absorbers, they are suited in particular when certain unwanted resonances are to be lowered. Three resonances are found in the frequency range up to 400 Hz. These are peaks at 60, 198 and 379 Hz. The resonance at 60 Hz corresponds to a rigid body mode similar to the rigid body mode that occurs at 0 Hz in a closed fluid-filled box, cf. Marburg et al. (2006). The resonances at 198 and 379 Hz correspond to half

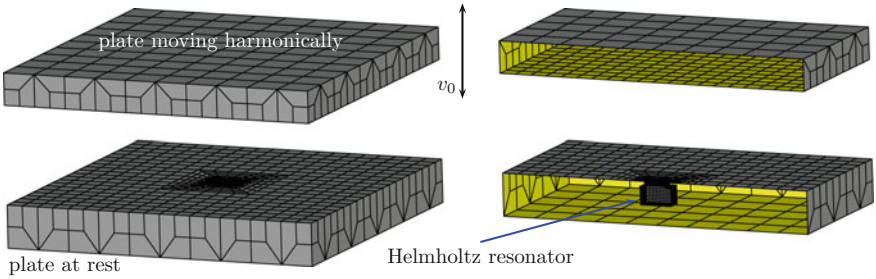


Fig. 38 Configuration of two parallel plates with a Helmholtz resonator designed into the *lower* fixed plate. The *upper* plate oscillates as a rigid piston

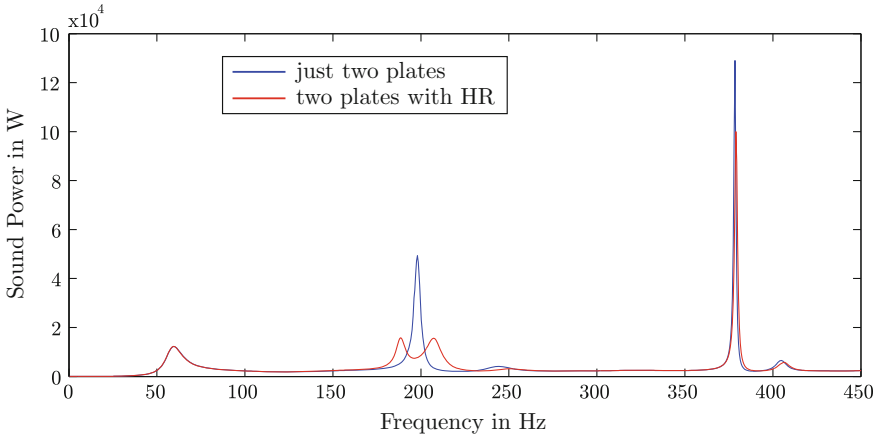


Fig. 39 Radiated sound power from the two plates with resonance peaks. The second resonance peak at 198 Hz is clearly lowered by the Helmholtz resonator in the lower plate

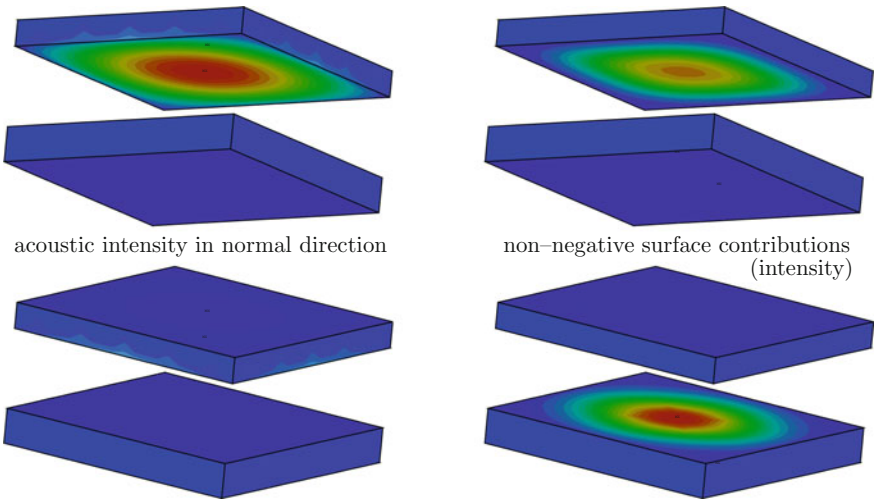


Fig. 40 Comparison between acoustic intensity in normal direction and non-negative surface contributions (intensity) for 198 Hz of configuration without Helmholtz resonator

a wavelength and one full wavelength between the plates, respectively. Adding a Helmholtz resonator tuned to a resonant frequency of 198 Hz to the lower plate of the open resonator significantly reduces the sound power at this frequency, while other frequencies remain mostly unaffected. This is the typical behavior of a tuned vibration absorber or Helmholtz resonator. Figure 39 shows the two curves of the radiated sound power.

Figures 40 and 41 compare the normal sound intensity and surface contribution for the two plate configurations (with and without a Helmholtz resonator). The normal

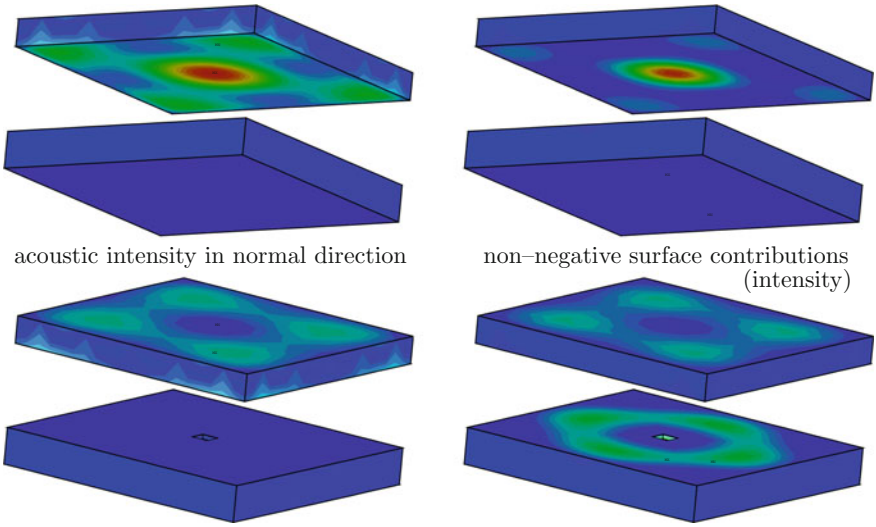


Fig. 41 Comparison between acoustic intensity in normal direction and non-negative surface contributions (intensity) for 198 Hz of configuration with Helmholtz resonator

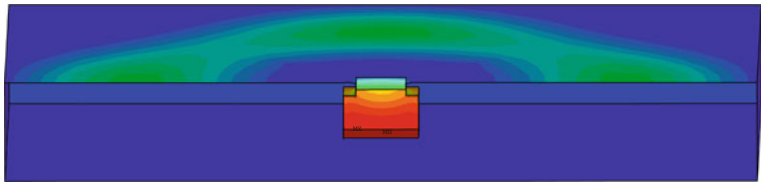


Fig. 42 Section view of Fig. 41 to illuminate the substantial non-negative surface contributions in the Helmholtz resonator at 198 Hz

sound intensity is always zero for the lower plate because of zero particle velocity. On the upper plate, the normal sound intensity is positive on the inner side facing the lower plate and negative on the outer side. In contrast, the surface contribution is distributed over both plates and is always positive.

It is important to note that the lower plate contributes to the radiated sound despite being fixed in space. The fact that the lower plate contributes to the radiated sound becomes obvious if the plate was removed from the vibroacoustic system in which case the frequency response of the system would change significantly.

The localized effect in the results for the surface contribution of the Helmholtz resonator on the fixed bottom plate can be clearly observed in the section view in Fig. 42. Thus, the surface contribution is more appropriate for visualization of the actual contributions of the lower and upper plates to the radiated sound power.

12 Conclusion

This chapter presented many details of the three dimensional boundary element method for the Helmholtz equation. It mainly focused on collocation methods and illuminated a number of other topics related to this. Among them, the use of discontinuous Lagrangian elements and the application of the Burton and Miller method for exterior problems have been recommended. The chapter discussed the motivation and the basic ideas of fast boundary element techniques, as suitable strategy for structure-fluid interaction and a formulation to consider symmetric and periodic models. Finally, a brief introduction into panel contribution analysis was provided.

While having discussed numerous issues of the boundary element method, such a survey can never be complete. An incomplete list of topics which have not been touched at all may read as follows:

- Efficient analysis over frequency ranges: Many technical problems require frequency range solutions. Only a few approaches appeared in literature on this. One of them is known as Padé and Padé-via-Lanczos approximation, cf. Coyette et al. (1999), Baumgart et al. (2007).
- Setup and solution of the eigenvalue problem: The implicitly frequency dependent matrices are not giving direct access to the acoustic eigenfrequencies. Approaches for this were presented in Chen et al. (1993), Zheng et al. (2015).
- Numerical damping and pollution effect: It is hardly known that the acoustic boundary element method produces numerical damping, cf. Marburg (2016). In the future, it will be shown by the author that this can be understood as a pollution effect which is well-known from the finite element method for wave phenomena.
- Thin radiators: In literature, radiation from thin plates and shells is often referred to as indirect BEM. Although common, the author is not convinced that this is a reasonable categorization. A nice approach which accommodates direct and indirect BEM under the roof of a Galerkin method was presented in Chen et al. (1997, 1998).
- Sound propagation above an impedance/admittance plane: Environmental acoustics often requires sound propagation above non-ideal halfplanes. This has been discussed (even together with moving sources) in Ochmann and Brick (2008), Ochmann (2013).
- Efficient sensitivity analysis: There are numerous approaches in literature including some contributions by the author, cf. Marburg (2002a), Fritze et al. (2005), Chen et al. (2016, 2017).
- Aeroacoustic analogies and their combination with BEM: Aeroacoustic analogies have many similarities with BEM and in some cases they were even efficiently coupled, cf. Croaker et al. (2013, 2015, 2016).

Overall, the author wishes to conclude that the boundary element method is a suitable tool for linear acoustic analysis in frequency domain.

References

Adey, R. A., Niku, S. M., Baynham, J., & Burns, P. (1995). Predicting acoustic contributions and sensitivity. Application to vehicle structures. In C. A. Brebbia (Ed.), Computational acoustics and its environmental applications (pp. 181–188). Southampton: Computational Mechanics Publications.

Atkinson, K. E. (1997). *The numerical solution of integral equations of the second kind* (1st ed.). Cambridge: Cambridge University Press.

Baumgart, J., Marburg, S., & Schneider, S. (2007). Efficient sound power computation of open structures with infinite/finite elements and by means of the Padé-via-Lanczos algorithm. *Journal of Computational Acoustics*, *15*, 557–577.

Bebendorf, M. (2008). *Hierarchical matrices: A means to efficiently solve elliptic boundary value problems*. Berlin: Springer.

Bebendorf, M., Kuske, C., & Venn, R. (2015). Wideband nested cross approximation for Helmholtz problems. *Numerische Mathematik*, *130*, 1–34.

Bernardi, C., Maday, Y., & Patera, A. T. (1994). A new nonconforming approach to domain decomposition: The mortar element method. In H. Brezis and J.-L. Lions(Eds.), Nonlinear partial differential equations and their applications (Vol. 11, pp. 13–51). Pitman, New York: College de France Seminar.

Bespalov, A. (2000). On the usage of a regular grid for implementation of boundary integral methods for wave problems. *Russian Journal of Numerical Analysis and Mathematical Modelling*, *15*, 469–488.

Bork, I. (2000). A comparison of room simulation software - the 2nd round robin on room acoustical computer simulation. *Acta Acustica united with Acustica*, *86*, 943–956.

Bork, I. (2005a). Report on the 3rd round robin on room acoustical computer simulation - Part I: Measurements. *Acta Acustica united with Acustica*, *91*, 740–752.

Bork, I. (2005b). Report on the 3rd round robin on room acoustical computer simulation - Part II: Calculations. *Acta Acustica united with Acustica*, *91*, 753–763.

Brakhage, H., & Werner, P. (1965). Über das Dirichlet'sche Außenraumproblem für die Helmholtz'sche Schwingungsgleichung. *Archiv der Mathematik*, *16*, 325–329.

Brancati, A., Aliabadi, M., & Milazzo, A. (2011). An improved hierarchical ACA technique for sound absorbent materials. *Computer Modeling in Engineering and Sciences*, *78*, 1–24.

Brebbia, C. A., Telles, J. F. C., & Wrobel, L. C. (1984). *Boundary element techniques*. Berlin: Springer.

Brunner, D., Junge, M., Rapp, P., Bebendorf, M., & Gaul, L. (2010). Comparison of the fast multipole method with hierarchical matrices for the Helmholtz-BEM. *Computer Modeling in Engineering and Sciences*, *58*, 131–160.

Burton, A. J., & Miller, G. F. (1971). The application of integral equation methods to the numerical solution of some exterior boundary-value problems. *Proceedings of the Royal Society of London*, *323*, 201–220.

Chen, P. T., & Ginsberg, J. H. (1995). Complex power, reciprocity, and radiation modes for submerged bodies. *Journal of the Acoustical Society of America*, *98*, 3343–3351.

Chen, Z. S., Hofstetter, G., & Mang, H. A. (1993). A 3D boundary element method for determination of acoustic eigenfrequencies considering admittance boundary conditions. *Journal of Computational Acoustics*, *1*, 455–468.

Chen, Z. S., Hofstetter, G., & Mang, H. A. (1997). A symmetric Galerkin formulation of the boundary element method for acoustic radiation and scattering. *Journal of Computational Acoustics*, *5*, 219–241.

Chen, Z. S., Hofstetter, G., & Mang, H. A. (1998). A Galerkin-type BE-FE formulation for elasto-acoustic coupling. *Computer Methods in Applied Mechanics and Engineering*, *152*, 147–155.

Chen, Z. S., Hofstetter, G., & Mang, H. (2008). A Galerkin-type be-formulation for acoustic radiation and scattering of structures with arbitrary shape. In S. Marburg & B. Nolte (Eds.), *Com-*

putational acoustics of noise propagation in fluids. Finite and boundary element methods (pp. 435–458). Berlin: Springer.

Chen, S., Liu, Y., & Dou, X. (2000). A unified boundary element method for the analysis of sound and shell-like structure interactions. II. Efficient solution techniques. *Journal of the Acoustical Society of America*, *108*, 2738–2745.

Chen, L., Chen, H., Zheng, C., & Marburg, S. (2016). Structural-acoustic sensitivity analysis of radiated sound power using a finite element/ discontinuous fast multipole boundary element scheme. *International Journal for Numerical Methods in Fluids*, *82*, 858–878.

Chen, L., Marburg, S., Chen, H., Zhang, H., & Gao, H. (2017). An adjoint operator approach for sensitivity analysis of radiated sound power in fully coupled structural-acoustic systems. *Journal of Computational Acoustics*, *25*, 1750003 (24 p.).

Ciskowski, R. D., & Brebbia, C. A. (Eds.). (1991). *Boundary elements in acoustics*. Southampton, Boston: Computational Mechanics Publications and Elsevier Applied Science.

Coyette, J.-P., Wynendaele, H., & Chargin, M. K. (1993). A global acoustic sensitivity tool for improving structural design. *Proceedings- SPIE The International Society for Optical Engineering*, *1923*, 1389–1394.

Coyette, J.-P., Lecomte, C., Migeot, J.-L., Blanche, J., Rochette, M., & Mirkovic, G. (1999). Calculation of vibro-acoustic frequency response functions using a single frequency boundary element solution and a Padé expansion. *Acustica*, *85*, 371–377.

Cremers, L., Guisset, P., Meulewaeter, L., & Tournour, M. (2000). A computer–aided engineering method for predicting the acoustic signature of vibrating structures using discrete models. Great Britain Patent No. GB 2000–16259.

Croaker, P., Kessissoglou, N., & Marburg, S. (2015). Strongly singular and hypersingular volume integrals for near-field aeroacoustics. *International Journal for Numerical Methods in Fluids*, *77*, 274–318.

Croaker, P., Kessissoglou, N. J., & Marburg, S. (2016). Aeroacoustic scattering using a particle accelerated computational fluid dynamics/boundary element technique. *AIAA Journal*, *54*, 2116–2133.

Croaker, P., Marburg, S., Kinns, R., & Kessissoglou, N. J. (2013). A fast low-storage method for evaluating Lighthill's volume quadrupoles. *AIAA Journal*, *51*, 867–884.

do Rego Silva, J. J. (1993). *Acoustic and elastic wave scattering using boundary elements* Topics in engineering (Vol. 18). Southampton, Boston: Computational Mechanics Publications.

Dong, J., Choi, K. K., & Kim, N.-H. (2004). Design optimization of structural-acoustic problems using FEA-BEA with adjoint variable method. *ASME Journal of Mechanical Design*, *126*, 527–533.

Fard, S. M. B., Peters, H., Kessissoglou, N., & Marburg, S. (2015). Three dimensional analysis of a noise barrier using a quasi-periodic boundary element method. *Journal of the Acoustical Society of America*, *137*, 3107–3114.

Fard, S. M. B., Peters, H., Marburg, S., & Kessissoglou, N. (2017). Acoustic performance of a barrier embedded with Helmholtz resonators using a quasi-periodic boundary element technique. *Acta Acustica united with Acustica*, *103*, 444–450.

Fernandez-Grande, E., Jacobsen, F., & Leclére, Q. (2012). Direct formulation of the supersonic acoustic intensity in space domain. *Journal of the Acoustical Society of America*, *131*, 186–193.

Flemisch, B., Kaltenbacher, M., & Wohlmuth, B. I. (2006). Elasto-acoustic and acoustic-acoustic coupling on nonmatching grids. *International Journal for Numerical Methods in Engineering*, *67*, 1791–1810.

Fritze, D., Marburg, S., & Hardtke, H.-J. (2005). FEM-BEM-coupling and structural-acoustic sensitivity analysis for shell geometries. *Computers and Structures*, *83*, 143–154.

Fritze, D., Marburg, S., & Hardtke, H.-J. (2009). Estimation of radiated sound power: A case study on common approximation methods. *Acta Acustica united with Acustica*, *95*, 833–842.

Fuß, S., Hawkins, S. C., & Marburg, S. (2011). An eigenvalue search algorithm for modal analysis of a resonator in free space. *Journal of Computational Acoustics*, *19*, 95–109.

Galkowski, J., Müller, E. H., & Spence, E. A. (2016). Wavenumber–explicit analysis for the Helmholtz h–BEM: error estimates and iteration counts for the Dirichlet problem. Preprint in numerical analysis, Cornell University. https://arxiv.org/abs/1608.01035.

Gaul, L., Brunner, D., & Junge, M. (2008). Coupling a fast boundary element method with a finite element formulation for fluid-structure interaction. In S. Marburg & B. Nolte (Eds.), Computational acoustics of noise propagation in fluids. Finite and boundary element methods (pp. 519–546). Berlin: Springer.

Giebermann, K. (2001). Multilevel representations of boundary integral operators. *Computing*, *67*, 183–207.

Gumerov, N. A., & Duraiswami, R. (2004). *Fast multipole methods for the Helmholtz equation in three dimensions*. Oxford: Elsevier Science & Technology.

Gumerov, N. A., & Duraiswami, R. (2009). A broadband fast multipole accelerated boundary element method for the three dimensional Helmholtz equation. *Journal of the Acoustical Society of America*, *125*, 191–205.

Harari, I., & Hughes, T. J. R. (1992). A cost comparison of boundary element and finite element methods for problems of time-harmonic acoustics. *Computer Methods in Applied Mechanics and Engineering*, *97*, 77–102.

Hornikx, M., Kaltenbacher, M., & Marburg, S. (2015). A platform for benchmark cases in computational acoustics. *Acta Acustica united with Acustica*, *101*, 811–820.

Ihlenburg, F. (1998). Finite element analysis of acoustic scattering (Vol. 132). Applied mathematical sciences. Berlin: Springer.

Ishiyama, S.-I., Imai, M., Maruyama, S.-I., Ido, H., Sugiura, N., & Suzuki, S. (1988). The application of ACOUST/BOOM - a noise level prediction and reduction code. *SAE-Paper*, *880910*, 195–205.

Jean, P., & Defrance, J. (2015). Sound propagation in rows of cylinders of infinite extent: Application to sonic crystals and thickets along roads. *Acta Acustica united with Acustica*, *101*, 474–483.

Karimi, M., Croaker, P., & Kessissoglou, N. (2016). Boundary element solution for periodic acoustic problems. *Journal of Sound and Vibration*, *360*, 129–139.

Karimi, M., Croaker, P., & Kessissoglou, N. (2017). Acoustic scattering for 3D multi-directional periodic structures using the boundary element method. *Journal of the Acoustical Society of America*, *141*, 313–323.

Kirkup, S. M. (1998). *The boundary element method in acoustics*. Heptonstall: Integrated Sound Software.

Koopmann, G. H., & Fahnline, J. B. (1997). *Designing quiet structures: A sound power minimization approach*. San Diego: Academic Press.

Kupradze, V. D. (1956). *Randwertaufgaben der Schwingungstheorie und Integralgleichungen*. Berlin: Deutscher Verlag der Wissenschaften. (1. Russian edition 1950).

Kussmaul, R. (1969). Ein numerisches Verfahren zur Lösung des Neumannschen Außenraumproblems für die Helmholtzsche Schwingungsgleichung. *Computing*, *4*, 246–273.

Lam, Y. W. (1999). A boundary integral formulation for the prediction of acoustic scattering from periodic structures. *Journal of the Acoustical Society of America*, *105*, 762–769.

Liu, Y. (2009). *Fast multipole boundary element method. Theory and applications in engineering*. New York: Cambridge University Press.

Liu, X., Wu, H., & Jiang, W. (2017). Hybrid approximation hierarchical boundary element methods for acoustic problem. *Journal of Computational Acoustics* (in print).

Liu, D., Peters, H., Marburg, S., & Kessissoglou, N. J. (2016a). Supersonic intensity and non-negative intensity for prediction of radiated sound. *Journal of the Acoustical Society of America*, *139*, 2797–2806.

Liu, D., Peters, H., Marburg, S., & Kessissoglou, N. J. (2016b). Surface contributions to scattered sound power using non-negative intensity. *Journal of the Acoustical Society of America*, *140*, 1206–1217.

Magalhaes, M. B. S., & Tenenbaum, R. A. (2006). Supersonic acoustic intensity for arbitrarily shaped sources. *Acta Acustica united with Acustica*, *92*, 189–201.

Marburg, S. (2002a). Developments in structural–acoustic optimization for passive noise control. *Archives of Computational Methods in Engineering. State of the Art Reviews*, *9*, 291–370.
Marburg, S. (2002b). Six boundary elements per wavelength. Is that enough? *Journal of Computational Acoustics*, *10*, 25–51.
Marburg, S. (2005). Normal modes in external acoustics. Part I. Investigation of the one-dimensional duct problem. *Acta Acustica united with Acustica*, *91*, 1063–1078.
Marburg, S. (2006). Normal modes in external acoustics. Part III: Sound power evaluation based on frequency-independent superposition of modes. *Acta Acustica united with Acustica*, *92*, 296–311.
Marburg, S. (2008). Discretization requirements: How many elements per wavelength are necessary? In S. Marburg & B. Nolte (Eds.), *Computational acoustics of noise propagation in fluids. Finite and boundary element methods* (pp. 309–332). Berlin: Springer.
Marburg, S. (2016a). The Burton and Miller method: Unlocking another mystery of its coupling parameter. *Journal of Computational Acoustics*, *24*, 1550016 (20 p.).
Marburg, S. (2016b). Numerical damping in the acoustic boundary element method. *Acta Acustica united with Acustica*, *102*, 415–418.
Marburg, S., & Amini, S. (2005). Cat's eye radiation with boundary elements: Comparative study on treatment of irregular frequencies. *Journal of Computational Acoustics*, *13*, 21–45.
Marburg, S., & Anderssohn, R. (2011). Fluid structure interaction and admittance boundary conditions: Setup of an analytical example. *Journal of Computational Acoustics*, *19*, 63–74.
Marburg, S., & Hardtke, H.-J. (1999). A study on the acoustic boundary admittance. Determination, results and consequences. *Engineering Analysis with Boundary Elements*, *23*, 737–744.
Marburg, S., & Hardtke, H.-J. (2003). Investigation and optimization of a spare wheel well to reduce vehicle interior noise. *Journal of Computational Acoustics*, *11*, 425–449.
Marburg, S., & Nolte, B. (Eds.). (2008a). *Computational acoustics of noise propagation in fluids. Finite and boundary element methods*. Berlin: Springer.
Marburg, S., & Nolte, B. (2008b). A unified approach to finite and boundary element discretization in linear time–harmonic acoustics. In S. Marburg & B. Nolte (Eds.), *Computational acoustics of noise propagation in fluids. Finite and boundary element methods* (pp. 1–34). Berlin: Springer.
Marburg, S., & Schneider, S. (2003a). Performance of iterative solvers for acoustic problems. Part I. Solvers and effect of diagonal preconditioning. *Engineering Analysis with Boundary Elements*, *27*, 727–750.
Marburg, S., & Schneider, S. (2003b). Influence of element types on numeric error for acoustic boundary elements. *Journal of Computational Acoustics*, *11*, 363–386.
Marburg, S., & Wu, T. W. (2008). Treating the phenomenon of irregular frequencies. In S. Marburg & B. Nolte (Eds.), *Computational acoustics of noise propagation in fluids. Finite and boundary element methods* (pp. 411–434). Berlin: Springer.
Marburg, S., Shepherd, M., & Hambric, S. A. (2016). Structural acoustic optimization. In S. A. Hambric, S. H. Sung, & D. J. Nefske (Eds.), *Engineering vibroacoustic analysis: methods and applications* (pp. 268–304). Chichester: Wiley.
Marburg, S., Hardtke, H.-J., Schmidt, R., & Pawandenat, D. (1997). An application of the concept of acoustic influence coefficients for the optimization of a vehicle roof. *Engineering Analysis with Boundary Elements*, *20*, 305–310.
Marburg, S., Rennert, R., Schneider, S., & Hardtke, H.-J. (2002). Resonances in external acoustics? An example of tire noise excitation. In A. Calvo-Manzano, A. Perez-Lopez, & J. S. Santiago (Eds.), *Proceedings of Forum Acusticum, Special Issue of Revista de Acustica* (Vol. 33, pp. 3–4). Sevilla, (CD).
Marburg, S., Schneider, S., Vorländer, M., & Romanenko, G. (2003). Boundary elements for room acoustic measures. In *Proceedings of the INTERNOISE 2003, The 32nd International Congress and Exposition on Noise Control Engineering, Held in Seogwipo/Korea* (pp. 3598–3604). Seoul: Covan International Corp.
Marburg, S., Dienerowitz, F., Horst, T., & Schneider, S. (2006). Normal modes in external acoustics. Part II: Eigenvalues and eigenvectors in 2D. *Acta Acustica united with Acustica*, *92*, 97–111.

Marburg, S., Lösche, E., Peters, H., & Kessissoglou, N. J. (2013). Surface contributions to radiated sound power. *Journal of the Acoustical Society of America*, *133*, 3700–3705.

Messner, M., Schanz, M., & Darve, E. (2012). Fast directional multilevel summation for oscillatory kernels based on Chebyshev interpolation. *Journal of Computational Physics*, *231*, 1175–1196.

Meyer, W. L., Bell, W. A., Zinn, B. T., & Stallybrass, M. P. (1978). Boundary integral solutions of three dimensional acoustic radiation problems. *Journal of Sound and Vibration*, *59*, 245–262.

Moheit, L., & Marburg, S. (2017). Infinite elements and their influence on normal and radiation modes in exterior acoustics. *Journal of Computational Acoustics*, *25*, 1650020 (20 p.).

Ochmann, M. (2013). Exact solutions for sound radiation from a moving monopole above an impedance plane. *Journal of the Acoustical Society of America*, *133*, 1911–1921.

Ochmann, M., & Brick, H. (2008). Acoustical radiation and scattering above an impedance plane. In S. Marburg & B. Nolte (Eds.), *Computational acoustics of noise propagation in fluids. Finite and boundary element methods* (pp. 459–494). Berlin: Springer.

Panič, O. I. (1965). K voprosu o razrešimosti vnešnich kraevich zadač dlja volnovogo uravnenija i dlja sistemi uravnenij Maxwella. *Uspechi Mathematiceskich Nauk*, *20*, 221–226.

Peters, H., Kessissoglou, N. J., & Marburg, S. (2012a). Enforcing reciprocity in numerical analysis of acoustic radiation modes and sound power evaluation. *Journal of Computational Acoustics*, *20*, 1250005 (19 p.).

Peters, H., Marburg, S., & Kessissoglou, N. J. (2012b). Structural-acoustic coupling on non-conforming meshes with quadratic shape functions. *International Journal for Numerical Methods in Engineering*, *91*, 27–38.

Peters, H., Kessissoglou, N. J., Lösche, E., & Marburg, S. (2013a). Prediction of radiated sound power from vibrating structures using the surface contribution method. In T. McMinn (Ed.), *Proceedings of Acoustics 2013 Victor Harbor: Science, Technology and Amenity. Proceedings of the Annual Conference of the Australian Acoustical Society*. (CD).

Peters, H., Kessissoglou, N. J., & Marburg, S. (2013b). Modal decomposition of exterior acoustic-structure interaction. *Journal of the Acoustical Society of America*, *133*, 2668–2677.

Peters, H., Kessissoglou, N., & Marburg, S. (2014). Modal decomposition of exterior acoustic-structure interaction problems with model order reduction. *Journal of the Acoustical Society of America*, *135*, 2706–2717.

Roozen-Kroon, P. J. M. (1992). *Structural optimization of bells*. Dissertation, Technische Universiteit Eindhoven.

Saad, Y., & Schultz, M. H. (1986). GMRES: A generalized minimal residual algorithm for solving non-symmetric linear systems. *SIAM Journal of Scientific and Statistical Computing*, *7*, 856–869.

Sakuma, T., Schneider, S., & Yasuda, Y. (2008). Fast solution methods. In S. Marburg & B. Nolte, (Eds.), *Computational acoustics of noise propagation in fluids. Finite and boundary element methods* (pp. 333–368). Berlin: Springer.

Schenck, H. A. (1968). Improved integral formulation for acoustic radiation problems. *Journal of the Acoustical Society of America*, *44*, 41–58.

Schneider, S. (2003). Application of fast methods for acoustic scattering and radiation problems. *Journal of Computational Acoustics*, *11*, 387–401.

Schneider, S. & Marburg, S. (2003). Performance of iterative solvers for acoustic problems. Part ii. Acceleration by ILU-type preconditioner. *Engineering Analysis with Boundary Elements*, *27*, 751–757.

Telles, J. C. F. (1987). A self-adaptive coordinate transformation for efficient numerical evaluation of general boundary element integrals. *International Journal for Numerical Methods in Engineering*, *24*, 959–973.

Thompson, L. L., & Pinsky, P. M. (1994). Complex wavenumber Fourier analysis of the p-version finite element method. *Computational Mechanics*, *13*, 255–275.

von Estorff, O. (Ed.). (2000). *Boundary elements in acoustics: advances and applications*. Southampton: WIT Press.

Vorländer, M. (1989). Simulation of the transient and steady state sound propagation in rooms using a new combined sound particle-image source algorithm. *Journal of the Acoustical Society of America*, *86*, 172–178.

Weyl, H. (1952). Kapazität von Strahlungsfeldern. *Mathematische Zeitschrift*, *55*, 187–198.

Wilkes, D. R., & Duncan, A. J. (2015). Acoustic coupled fluid-structure interactions using a unified fast multipole boundary element method. *Journal of the Acoustical Society of America*, *137*, 2158–2167.

Williams, E. G. (1995). Supersonic acoustic intensity. *Journal of the Acoustical Society of America*, *97*, 121–127.

Williams, E. G. (2013). Convolution formulations for non-negative intensity. *Journal of the Acoustical Society of America*, *134*, 1055–1066.

Wu, T. W. (2000a). The Helmholtz integral equation. In T. W. Wu (Ed.), *Boundary element in acoustics: Fundamentals and computer codes* (pp. 9–28). Southampton: WIT Press.

Wu, T. W. (Ed.). (2000b). *Boundary element acoustics: fundamentals and computer codes*. Southampton: WIT Press.

Wu, S. F., & Natarajan, L. K. (2013). Panel acoustic contribution analysis. *Journal of the Acoustical Society of America*, *133*, 799–809.

Wu, T. W.,& Seybert, A. F. (1991). Acoustic radiation and scattering. In R. D. Ciskowski & C. A. Brebbia (Eds.), *Boundary elements in acoustics* (pp. 61–76). Southampton: Computational Mechanics Publications; London: Elsevier Applied Science.

Wu, H. J., Liu, Y. J., & Jiang, W. K. (2013). A low-frequency fast multipole boundary element method based on analytical integration of the hypersingular integral for 3D acoustic problems. *Engineering Analysis with Boundary Elements*, *37*, 309–318.

Wu, S. F., Moondra, M., & Beniwal, R. (2015). Analyzing panel acoustic contributions toward the sound field inside the passenger compartment of a full-size automobile. *Journal of the Acoustical Society of America*, *137*, 2101–2112.

Zheng, C.-J., Chen, H.-B., Gao, H.-F., & Du, L. (2015). Is the Burton-Miller formulation really free of fictitious eigenfrequencies? *Engineering Analysis with Boundary Elements*, *59*, 43–51.

Ziegelwanger, H., Reiter, P., & Conter, M. (2017). The three-dimensional quasi-periodic boundary element method: Implementation, evaluation, and use cases. *International Journal of Computational Methods and Experimental Measurements*, *5*, 404–414.

Direct Aeroacoustic Simulations Based on High Order Discontinuous Galerkin Schemes

Andrea Beck and Claus-Dieter Munz

Abstract In this chapter, we discuss some of the challenges that arise for the direct numerical computation of noise generation and transport. Noise sources are associated with the non-linearities of the underlying hydrodynamics, i.e. with the turbulent fluctuations across the energy spectrum. Thus, the numerical resolution of these sound sources not only inherits the numerical difficulties that arise for general DNS and LES of turbulent flows, but the scale separation between the hydrodynamic velocity fluctuations and the radiated pressure waves adds additional challenges, for example in terms of boundary conditions and numerical approximation accuracy. Therefore, a highly efficient and accurate numerical scheme is necessary. The framework presented herein is based on a particular version of the Discontinuous Galerkin method, in which a nodal as well as discretely orthogonal basis is used for computational efficiency. This discretization choice allows arbitrary order in space while also supporting unstructured meshes. After discussing the details of the framework, examples of direct noise computation are presented, with a special focus on the numerical simulation of acoustic feedback in a complex automotive application.

1 Introduction

The field of aeroacoustics encompasses the sound waves generated by and propagated through unsteady turbulent or vortical aerodynamic internal and external flows. Computational methods that are aimed at simulating these sound waves or their effects are often termed Computational Aeroacoustics (CAA) methods, which are a subset of the more general Computational Fluids Dynamics (CFD) field. The main challenge in the numerical approximation of aeroacoustics stems from its multiscale nature. The generating mechanism is inherently non-linear and unsteady, which precludes - without strong assumptions - time-averaged simulation schemes like the Reynolds-averaged Navier–Stokes (RANS) approach. Different theoretical

A. Beck (✉) · C.-D. Munz
Numerics Research Group, Institute of Aerodynamics and Gasdynamics,
Pfaffenwaldring 21, 70569 Stuttgart, Germany
e-mail: beck@iag.uni-stuttgart.de

M. Kaltenbacher (ed.), *Computational Acoustics*, CISM International Centre
for Mechanical Sciences 579, DOI 10.1007/978-3-319-59038-7_4

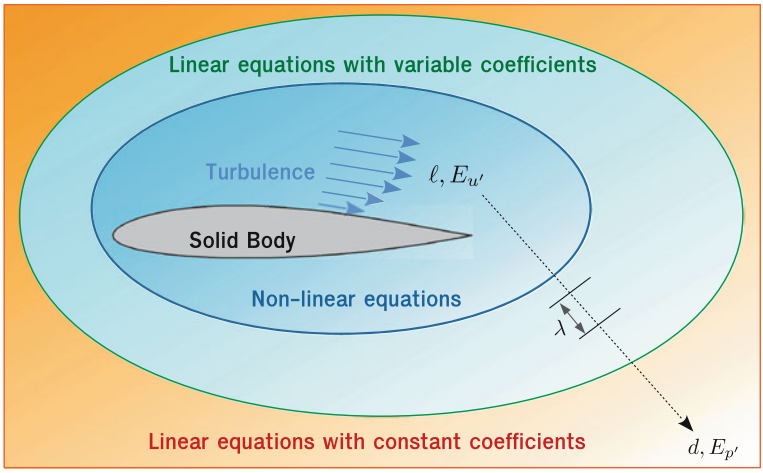


Fig. 1 Equations regimes and scales for aeroacoustics

interpretations regarding the source of sound generation exist, with the classical one focusing on the role of velocity fluctuations (Lighthill 1952) and a more recent one emphasizing vortical structures as the source mechanism for sound waves (Howe 2003). From both points of view however, it is clear that a successful simulation of aeroacoustics must include an accurate resolution of these generating mechanisms.

From a *physical* point of view, aeroacoustic problems can be categorized into those governed by non-linear effects, i.e. mainly the generation of pressure fluctuations from non-linear hydrodynamics, and those that are essentially linear, e.g. radiation into the far field, refraction or scattering. In Fig. 1, the relevant sets of equations (and thus the applicable numerical approaches discussed below) and the relevant scales are shown for a typical CAA problem.

Along a solid body, a turbulent boundary layer develops and radiates noise into its surroundings. It interacts with the trailing edge and thereby produces sound through the enhanced non-linearities in that region. A typical characteristic length of this generation mechanism is the boundary layer thickness ℓ. The energy content of the velocity fluctuations is characterized by $E_{u'}$. Since the relevant processes for the sound generation are inherently unsteady and non-linear, the full compressible Navier–Stokes equations are necessary to describe these processes accurately. As the time scale of the sound waves matches that of its source, the resulting wave length of the radiated sound is directly proportional to the speed of sound c, which explains the large discrepancy between λ and ℓ for low Mach number flows. For increasing Mach numbers, this clear scale separation vanishes. Further away from the solid body, the influence of non-linearities and viscosity is reduced and thus acoustic source terms vanish. Here, Euler equations and their linearized version (LEE) can be used to simulate acoustic transport by a background flow field, in which source terms generate the sound waves. Since the characteristic length ℓ now no longer needs to be resolved, the acoustic wave lengths λ and the associated wave speed

c now determine the spatial and temporal resolution requirements. This so-called hybrid approach, which is discussed in more detail below, thus explicitly exploits the scale separation. Even further away from the sound source, when $d/\lambda \gg 1$ and thus the source region becomes acoustically compact and reflections or diffraction are negligible, integral approaches based on the wave equation can be used to propagate the sound through the far field.

An example that clearly highlights the difficulties arising from the scale separation between the hydrodynamic source and the acoustic waves is given in a review by Lele: For a supersonic jet, the ratio of $E_{p'}/E_{u'}$ is only about 0.01, while for most other noise generation mechanisms, this ratio is even considerably smaller (Lele 1997).

From a *computational* point of view, the CAA methods (based on the equations discussed in Fig. 1) can be classified into two broad categories: The *direct* approach, labeled Direct Noise Computation (DNC) and the indirect or *hybrid* approach.

The direct approach is based on first principles and avoids any modeling approximations. It does not introduce an a priori conceptual split into a flow or hydrodynamic part and an acoustic part, but solves the full compressible flow equations which contain the sound generation mechanisms through non-linear vortical interactions. The solution to this single set of equations then contains all the physical effects included in the equations together with their coupling. In particular, acoustic feedback onto the flow field is naturally included in this approach.

For a meaningful DNC, numerical schemes must thus capture the unsteady solution over a wide range of local flow scales (to account for the non-linear source effects) and across large spatial and temporal scales (to account for the typically large length scales of the acoustic waves compared to the hydrodynamics). It therefore mandates numerical schemes that are capable of high local resolution as well as efficient and accurate long-term wave transport. While Direct Numerical Simulation (DNS), which resolves all flow scales including the dissipation range, is the preferred method of choice, it is essentially restricted by the direct dependency of its cost on the flow Reynolds number Re. Resolving only the dynamically important scales and modeling the isotropic parts by a subgrid scale model in a Large Eddy Simulation (LES) ameliorates this restriction and expands the range of applicability of DNC, but introduces the additional complexity of subgrid closure.

Another important conceptual complexity in DNC which should not be overlooked is the postprocessing of the solution, i.e. the a posteriori identification of acoustic sources and acoustic transport from the compressible flow field.

In the hybrid approach, the computation of the flow is decoupled from the computation of the sound. CAA then becomes a two-step, forward-coupled simulation approach. This separation is motivated by the disparity between the large length scales and low energy content of the acoustic field compared to the hydrodynamic field, i.e. the fundamental assumption states that while the unsteady vortical flow field generates sound waves and influences their propagation, these waves do not influence the flow field and act as a passive sink for the acoustic energy. This presumption is a good approximation in low Mach number flows, with the exception of acoustic feedback mechanisms.

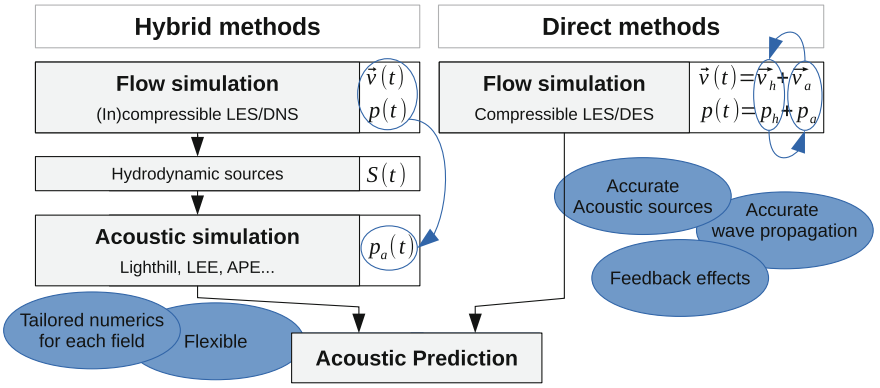


Fig. 2 Direct and hybrid acoustic simulation strategies, with permission from Frank (2016)

The decoupling of the two physical phenomena allows the development of numerical schemes tailored to the specific area of application. The governing equations, solution algorithms and discretizations can be chosen independently for each step to optimize their respective efficiency. For the hydrodynamic simulation, time-resolving simulation methods (DNS, LES, unsteady RANS) are used to compute a space-time evolution of the flow field from both the compressible and incompressible Navier–Stokes equations. From this solution, time- and space-dependent acoustic source terms for the subsequent acoustic simulation are generated. These source terms are then re-introduced in the second simulation step, for which different formulations for the propagation of acoustics exist. Many of these formulations are based on the inhomogeneous wave equation derived by Lighthill (1952) or a perturbation formulation of the Euler equations, and differ in their assumed relationship between the hydrodynamics and the source term and in the assumed state of the base flow. Beyond the flexibility in the choice of the equations and discretizations, the hybrid approach also allows the selection or truncation of the source term region and thus the isolation of different acoustic effects in the (computationally much cheaper) second step.

Besides the additional complexity stemming from the handling of the source term and the two computational schemes, the most serious drawback of the hybrid approach are its inherent underlying assumptions, which for example rule out acoustic feedback loop like the one discussed in Sect. 4.3. Also, if no clear scale separation exists (e.g. in high Mach number flows), the hybrid approach loses its validity and makes the designation of source terms ambiguous. Figure 2 summarizes the conceptual differences between direct and hybrid acoustic simulation strategies.

2 Numerical Schemes for Direct Acoustics

In this section, we will give a brief overview of the requirements and challenges for the scale resolving simulation of acoustics. Schemes used in direct methods, where the hydrodynamic and acoustic scales are resolved (see Figs. 1 and 2), clearly need good

scale-resolving capabilities. To a lesser extent, the requirements on the numerical scheme also apply to hybrid methods, in which hydrodynamic source computations and acoustic simulation are split into two subsequent simulations, which allow an independent discretization of each problem. Still, since the basic requirement is that of a multi-scale problem, essentially the same challenges to the discretization schemes exist.

2.1 Physical Considerations

As discussed in Sect. 1, the aeroacoustic sources stem from the hydrodynamic fluctuations, i.e. the frequency of the small scales of turbulent motion determines the bandwidth of the acoustics. From Kolmogorov's theory (1999), the relationship between the spatial scales in fully developed turbulence is known to be

$$\frac{L}{\eta} \sim Re^{3/4}, \tag{1}$$

i.e. the bandwidth between the largest or energy-carrying scales L and the smallest or dissipative scales η is determined by the Reynolds number. To estimate the local time scale associated with each wavenumber, an eddy-turnover-frequency (Colonius and Lele 2004) can be constructed on dimensional grounds from

$$f(k) = \frac{1}{\tau(k)} = \sqrt{k^3 E(k)}, \tag{2}$$

where $E(k)$ denotes the one-dimensional spectrum of kinetic energy. Using a von Kármán–Kraichnan model spectrum for $E(k)$, Fig. 3 depicts the eddy-turnover-frequency $f(k)$ for different Reynolds numbers. As a direct consequence of the increase in spatial bandwidth with Re, the range of f and its magnitude also increase, which results in a broader range of acoustic emission and shorter acoustic wavelengths. Thus, since the turbulent scales of motion and the noise generation mechanism are so closely coupled, the numerical simulation of noise generation is faced with the same issues as the scale-resolving simulation of (compressible) turbulence: The range of scales that can be resolved without the additional assumptions or models is limited by the wave propagation properties of the numerical scheme and its computational efficiency. In addition, in particular for low Mach number flows, the inefficient transfer from hydrodynamic to acoustic energy results in large discrepancies between the flow and acoustic energy, which makes the latter even more susceptible to approximation errors. One situation where LES can be applied successfully to acoustic problems without an explicit closure approach is when the dominant source mechanisms are associated with the 'large' flow scales, i.e. when the sound producing features of the flow are within in the range of scales that are well-resolved in an LES and essentially decoupled from the model errors. One example of such a situation will be given in Sect. 4.3.

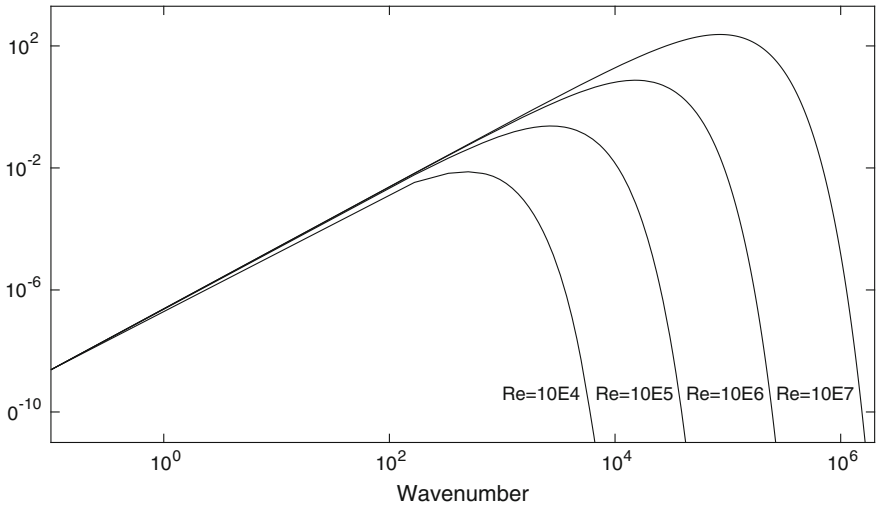


Fig. 3 Eddy turnover frequency for high Reynolds number flows

2.2 Discretization Methods for Scale Resolving Simulations

The requirements for LES and DNS discussed above and those for noise computation are of the same nature. They both imply that a numerical scheme of high or spectral order of accuracy is favorable, since these provide favorable wave resolution properties due to low approximation errors. A straightforward extension of Eq. (1) to three dimensions gives the following estimate for the number of degrees of freedom required for the spatial discretization operator of a DNS at a given Reynolds number

$$N_{3D} = \left(\frac{L}{\eta/n_{ppw}}\right)^3 \sim n_{ppw}^3 \, Re^{9/4}. \tag{3}$$

Here, n_{ppw} is the number of solution or grid points required to resolve structures of size η with a given approximation error. It thus can be interpreted as a number of points per wavelength criterion, which directly represents the numerical accuracy per degree of freedom. A more refined analysis leads to a more stringent requirement of $N_{3D} \sim n_{ppw}^3 \, Re^{37/14}$ (Choi and Moin 2012). Considering not only the spatial degrees of freedom, but also the fact that the characteristic time scale of the dissipation scales is directly proportional to η, the total computational cost in terms of spatial and temporal degrees of freedom N_{total} becomes

$$N_{total} \sim n_{ppw}^4 \, Re^3. \tag{4}$$

Clearly, not only the physical complexity of the problem can make or break a simulation through the dependence on Re, but also the numerical capabilities of the

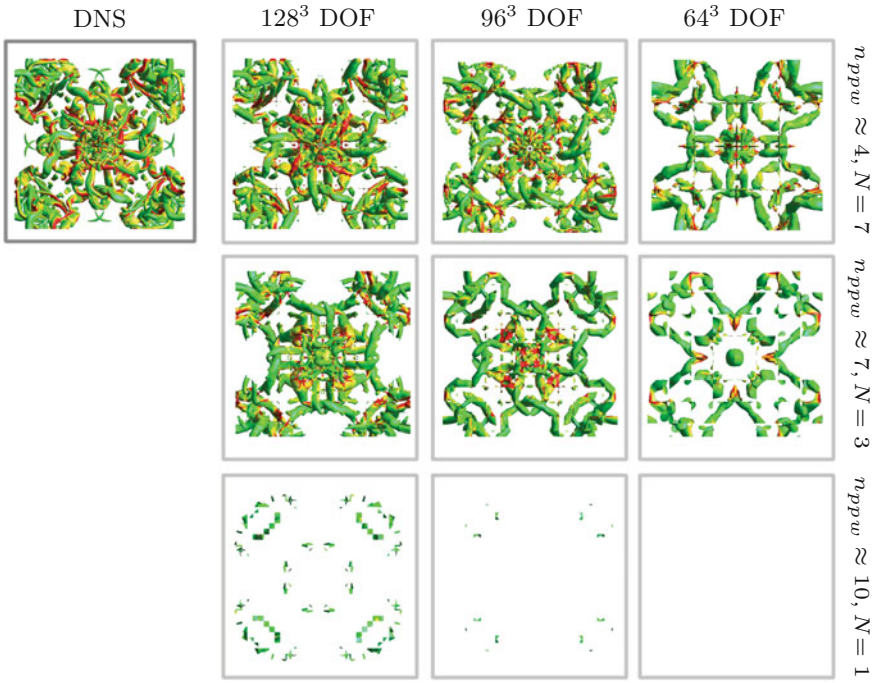


Fig. 4 Visualization of vortical structures of the Taylor–Green vortex at $Re = 1600$ via the λ_2 criterion for different discretizations

discretization scheme can be decisive. Thus, a fundamental demand for efficient numerical simulation of all or a subset of the scales of turbulent motion can be formulated as: The number of degrees of freedom or grid points required to accurately resolve the smallest occurring relevant scale, n_{ppw}, must be minimized. By their design, schemes of a high approximation order achieve this purpose for smooth problems. But also for under-resolved situations, these schemes can retain their low approximation errors over a wide range of resolved scales (Beck et al. 2014).

Figure 4 highlights the influence of the chosen discretization on the scale-resolving capabilities for turbulent flows. Shown is a visualization of the vortical structures of the Taylor–Green vortex at $Re = 1600$ and at non-dimensional time $t = 9$. The vortices are identified by the $\lambda_2 = -0.3$ criterion. Since the problem contains a number of symmetries, only one eighth of the full domain is shown. All computations are conducted with the Discontinuous Galerkin method presented in Sect. 3, where the polynomial degree N of the solution approximation and thus the order of accuracy can be chosen arbitrarily. In the upper left corner, the DNS result, computed with 512^3 DOF, is shown as a reference. In the 3×3 matrix to the right, each column corresponds to a fixed spatial number of DOF, and each row corresponds to a value of n_{ppw}. In other words, each row represents an h-refinement/coarsening of a given discretization, while each column shows different combinations of number of

elements and degree N. For example, in the column corresponding to 128^3 DOF, the first row entry is computed on a grid with 16^3 elements. In each element, the solution is approximated by a tensor product of one-dimensional polynomials of degree $N = 7$, leading to a total of 128^3 DOF. This high order approximation has a low value of $n_{ppw} \approx 4$ (Gassner and Kopriva 2011). In the second row, the number of elements is doubled per dimension to 32^3, while N is reduced to 3 with $n_{ppw} \approx 7$ leading again to 128^3 DOF.

Comparing the columns in Fig. 4, it is evident that as expected the solution quality deteriorates with respect to the DNS when the overall resolution is reduced. Grid artifacts become visible, and the small scale turbulent structures disappear, while the larger scale structures become smeared by the numerical diffusion. The more interesting observation from this plot comes from comparing the rows among each other. The n_{ppw} criterion clearly determines the scale resolving capabilities of the scheme, and for the same number of DOF, the solution produced by the second order scheme is completely dominated by the numerical errors.

From this discussion it follows that high order schemes are advantageous when considering acoustic sources and wave transport. However, high order accuracy and a low n_{ppw} is not the only determining factor for computational efficiency, but an important one. There are various ways in which discretizations achieve high order approximations, but they differ in other important aspects that overall determine their suitability for large scale direct noise computation. In the following, a brief overview of the typical discretization strategies is given.

For finite volume (FV) schemes, the integral form of the conservation equation is solved at a discrete level, i.e. the evolution of the mean in each grid cell is computed. Information exchange between the elements occurs via a numerical flux function. This ensures local conversation and introduces stability for underresolved problems and strong gradients. To achieve a higher order approximation, a reconstruction step is added, which reconstructs higher order approximation polynomials from the integral data across given element stencils. The specific methods then differ in the choice of the reconstruction stencils and in the combination or selection of the polynomials. In particular for three-dimensional simulations, this reconstruction process incurs a high computational effort and a complex parallelization. On non-regular grids, the formally high order accuracy is usually not obtained, which negates one main advantage of FV schemes, namely their general suitability for unstructured meshes. Figure 5 highlights the challenges introduced by the reconstruction process.

High order finite difference (FD) schemes are based on analytical differentiation of an interpolating polynomial. Thus, they inherit the simplicity of the interpolation operation, but also its drawbacks. For advection dominated problems, stencil upwinding or filtering is needed for stabilization. Achieving a high approximation order is straight-forward by stencil extension, but makes parallelization particularly demanding (alternatively, compact FD schemes solve a local linear system of equations). One subfamily of FD schemes are the dispersion relation preserving schemes, which sacrifice the theoretical order of convergence for improved phase and amplitude errors (Tam and Webb 1993; Bogey and Bailly 2004). The main drawback of FD

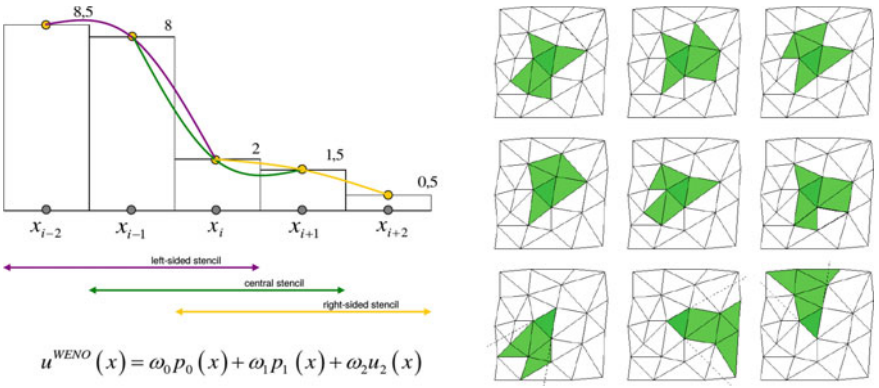


Fig. 5 *Left*: 1D stencil for quadratic reconstruction, *Right*: stencil choices for quadratic reconstruction in 2D

schemes is their reliance on structured grids and the complex integration of boundary conditions.

Global spectral (GS) methods form another class of schemes that have by design very favorable n_{ppw} values, down to the theoretical limit of 2 for Fourier-basis based methods. They have in common that the solution in the domain is approximated by a unique global solution representation, i.e. their stencil includes all available information. The residual is either minimized in an L_2-projection sense or at discrete solution points, leading to the Spectral-Galerkin-type schemes and the Spectral-Collocation-type schemes. These methods have been widely used in basic turbulence research, mainly for the incompressible Navier–Stokes equations (Yokokawa et al. 2002). For compressible problems, additional stabilization mechanism are required, e.g. Hussaini et al. (1985), Shebalin (1993). The global nature of the approximation makes parallelization non-trivial and costly compared to other methods. The main drawback of these methods is however their restriction to a single domain geometry.

In contrast to these global spectral methods, the class of high order finite element (FE) spectral methods decomposes the computational domain into grid cells or elements, which can be arranged in an unstructured, non-conforming way, akin to FV grids. Based on the chosen ansatz, this class can be split into continuous (for a globally continuous ansatz) and discontinuous (for an element-local ansatz) Galerkin methods. Both groups allow an easy way to increase the approximation order and thus reduce n_{ppw}. Continuous Galerkin methods are employed for incompressible flows mainly, and require additional stabilization for hyperbolic problems. Discontinuous Galerkin methods gain stability for compressible problems through the numerical flux function that penalizes inter-element discontinuities. In addition, the coupling through the fluxes and not the solution itself reduces the communication footprint of the method, and makes its parallelization straightforward. These methods thus combine high order accuracy, geometric flexibility and computational efficiency.

Table 1 Comparison of features of discretization schemes for direct acoustic computation

	n_{ppw}	Costs/DOF	Geometry	Parallelization	Stability
GS	✓	∼	–	∼	–
CG/DG	✓	✓	✓	✓	(✓)
FD	✓	✓	∼	✓	✓
FV HO	✓	∼	✓	∼	✓

Table 1 summarizes the advantages and disadvantages discussed here. For direct noise computation in complex domains, a single domain method is not practical so discretizations that rely on a global solution representation are ruled out. Furthermore, if geometric flexibility is required, only discretization strategies that naturally support unstructured meshing are viable options. Among these, DG methods combine high order accuracy without increasing stencil size and inherent suitability for hyperbolic problems, which make it a very suitable candidate as a base scheme for investigating noise generation. In the following section, we will present the numerical and implementation details of such a DG framework.

3 Discontinuous Galerkin Spectral Element Method

In this section, we present the details of a special variant of the DG method, namely the Discontinuous Galerkin Spectral Element Collocation Method (DGSEM). Discontinuous Galerkin methods in general can be interpreted as a hybrid of high order FE methods and FV methods, which gives them a number of favorable properties for scale-resolving simulations:

- Spectral accuracy for smooth problems when increasing the degree of the local ansatz (p-refinement), which results in low n_{ppw} requirements, as discussed in Sect. 2.2
- Natural support of arbitrarily shaped grid elements, which can be connected in an unstructured, non-conforming way
- Local grid refinement or basis enrichment in regions of interest (h/p-refinement)
- Stability for hyperbolic problems with discontinuities through numerical flux functions
- Local conservation for each element
- Weak imposition of the boundary conditions through fluxes
- Efficient parallelization due to minimal inter-element coupling
- Orthogonal hierarchical bases which resolve a large wave range within an element and which can be exploited in multiscale modeling

DG methods have a relatively recent research history. They were introduced by Reed and Hill (1973) in 1973 for linear advection problems of neutron transport on triangular meshes and analyzed by Lesaint and Raviart (1974). Research then lay dormant for about two decades, until Cockburn and Shu provided a systematic extension to systems of non-linear conversation equations (Cockburn and Shu 1991, 1989; Cockburn et al. 1989, 1990) such as e.g. the compressible gas dynamics. Bassi and Rebay were the first to introduce a mixed finite element type approach for the discontinuous Galerkin discretization of viscous flow problems and extended the DG method to the compressible Navier–Stokes equations (Bassi and Rebay 1997). Collis in 2002 was the first to use a high order DG method ($p = 6$) for the DNS of a weakly compressible turbulent channel flows at a low Reynolds number, with about 13 mio DOF for his finest mesh (Collis 2002).

Since DG methods are closely related to high order FE methods, the core of the method can be summarized in two steps: The projection operator of the variational formulation and the inversion of the mass matrix. The discretization and implementation choices for these two steps, together with the choice of the element topography, lead to different DG formulations. Among these (spatial) choices are the basis functions (e.g. Lagrange or Legendre-type polynomials), the approximation space spanned by these functions in multi-dimensions (a tensor-product approach or a full order basis), the choice of the quadrature method, the weak or strong DG-formulation, the discretization choices for the inviscid and viscous surface fluxes and the treatment of non-linearities. The temporal integration introduces another level of possible choices.

Among these different variants, the Discontinuous Galerkin Spectral Element Collocation Method (DGSEM) (Kopriva 2009; Hindenlang et al. 2012) combined with an explicit time integration scheme has shown to be highly effective and competitive for scale-resolving simulations.

3.1 Basic DG Discretization

In this section, we derive details and specific implementation choices of the Discontinuous Galerkin Spectral Element Collocation Method for a system of hyperbolic-parabolic conservation equations, following Kopriva (2009) and Hindenlang et al. (2012). Since the main focus is the direct noise computation, we use the compressible Navier–Stokes equations in physical space $\mathbb{R}^3$ as an example.

Compressible Navier–Stokes Equations The temporal and spatial evolution of a viscous, compressible fluid is governed by the conservation statements for mass, momentum and energy. In conservative form this set of partial differential equations for a Newtonian fluid is given as

$$\begin{aligned} \frac{\partial \rho}{\partial t} + \frac{\partial \left(\rho u_j\right)}{\partial x_j} &= 0, \\ \frac{\partial \left(\rho u_i\right)}{\partial t} + \frac{\partial \left(\rho u_i u_j + p\delta_{ij}\right)}{\partial x_j} &= \frac{\partial \sigma_{ij}}{\partial x_j}, \\ \frac{\partial \left(\rho e\right)}{\partial t} + \frac{\partial \left[\left(\rho e + p\right) u_j\right]}{\partial x_j} &= -\frac{\partial q_j}{\partial x_j} + \frac{\partial \left(\sigma_{ij} u_i\right)}{\partial x_j}. \end{aligned} \tag{5}$$

Here, the Einstein summation convention applies, δ_{ij} denotes the Kronecker delta function and $i, j = 1, 2, 3$. The conservative variables of mass, momentum and energy are $U = [\rho, \rho u_1, \rho u_2, \rho u_3, \rho e]$, where ρ denotes the density, u_i the ith component of the velocity vector and the total energy ρe is given by

$$\rho e = \rho \left(\frac{1}{2} u_i u_i + c_v T \right). \tag{6}$$

Herein, c_v and T denote the specific heat at constant volume and the temperature, respectively. The equation of a perfect gas is used to close the system:

$$p = \rho R T, \quad \gamma = \frac{c_p}{c_v}, \tag{7}$$

with $R = c_p - c_v$ as the specific gas constant, the pressure p and the adiabatic exponent γ. The viscous stress tensor σ_{ij} is a function of the viscosity μ (which itself is dependent on temperature) and the velocity gradient tensor

$$S_{ij} = \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} - \lambda \delta_{ij} \frac{\partial u_k}{\partial x_k}. \tag{8}$$

The bulk viscosity coefficient λ is commonly chosen to be $\frac{2}{3}$, which removes the trace from S_{ij}. The remaining unknown in Eq. (5) is the definition of the heat flux vector q_j as

$$q_j = -k \frac{\partial T}{\partial x_j}, \quad \text{with } k = \frac{c_p \mu}{Pr}, \tag{9}$$

where Pr denotes the Prandtl number of the fluid.

In vectorial form, Eq. (5) can be recast as

$$\begin{aligned} &\frac{\partial U}{\partial t} + \frac{\partial}{\partial x} F^c(U) + \frac{\partial}{\partial y} G^c(U) + \frac{\partial}{\partial z} H^c(U) \\ &\quad - \frac{\partial}{\partial x} F^v(U, \nabla_x U) - \frac{\partial}{\partial y} G^v(U, \nabla_x U) - \frac{\partial}{\partial z} H^v(U, \nabla_x U) = 0 \end{aligned} \tag{10}$$

with the vector of conserved variables U and the associated inviscid and viscous physical fluxes $\{F^c, G^c, H^c\}$ and $\{F^v, G^v, H^v\}$. Collecting the directional fluxes, this can further be simplified to the standard compact form of a conservation law

$$\begin{aligned} \frac{\partial U}{\partial t} + \nabla_x \cdot \vec{F}^c(U) - \nabla_x \cdot \vec{F}^v(U, \nabla_x U) = 0, \\ \frac{\partial U}{\partial t} + \nabla_x \cdot \vec{F}(U, \nabla_x U) = 0. \end{aligned} \tag{11}$$

Together with suitable initial and boundary conditions, Eq. (11) describes a system of conservation equation of hyperbolic-parabolic type, that can be now be discretized by the DGSEM method.

Spatial Discretization In order to solve this system of equations, a discretization of the computational domain consisting of non-overlapping elements is defined. In the DGSEM method, the type of elements is restricted to hexahedral cells which support a tensor product basis. The elements can be connected in a fully unstructured way. This restriction of the element type can be ameliorated by the use of non-conforming grids, but in general it makes the grid generation process more costly.

Once the grid has been created, each element in the physical domain is then mapped to a unit reference element $E \in [-1, 1]^3$ with coordinates $(\xi^1, \xi^2, \xi^3)^T$. The associated mapping function $\vec{x}(\vec{\xi})$ from reference to physical space is approximated as a polynomial itself and is then used to calculate the Jacobian $J(\vec{\xi}) = \det(\frac{\partial \vec{x}}{\partial \vec{\xi}})$. Clearly, for the mapping to be defined and invertible, $J(\vec{\xi})$ has to be positive everywhere, which can be challenging for non-linear mappings of curved elements (Hindenlang 2014). The main reason for the mapping step is to be able to define the operator itself in reference space, which means that a single shared set of basis functions and quadrature coefficients for each element exists.

The resulting individual element-based mapping is then used to transform Eq. (11) to reference space

$$U_t + \frac{1}{J(\vec{\xi})} \nabla_\xi \cdot \vec{\mathcal{F}}(U, \nabla_x U) = U_t + \frac{1}{J(\vec{\xi})} \nabla_\xi \cdot (\vec{\mathcal{G}}(U) - \vec{\mathcal{H}}(U, \nabla_x U)) = 0, \tag{12}$$

where $J(\vec{\xi}) := \vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)$ is again the Jacobian of the mapping $\vec{x}(\vec{\xi})$, calculated from the covariant basis vectors $\vec{a}_l := \frac{\partial \vec{x}}{\partial \xi^l}$. The covariant transformed fluxes are given by

$$\mathcal{F}^l := J\vec{a}^l \cdot \vec{F}, \quad l = 1, 2, 3, \tag{13}$$

with the metric terms

$$J\vec{a}^l := \vec{a}_k \times \vec{a}_m \quad (l, k, m) \text{ cyclic}. \tag{14}$$

The way the metric terms are discretized and implemented is important for the free-stream preserving property of the resulting method. We refer to Kopriva (2006) for

a discussion. This property ensures that the divergence operator remains zero for a spatially constant flux on a discrete level. Besides conservation, this property is of particular importance for acoustic propagation, where the energy of the acoustic waves is considerably smaller than that of the hydrodynamics and can easily be overwhelmed by small scale error terms.

Since the equation for each element is now defined in a common reference element, a shared polynomial basis can now be chosen. In DGSEM, the solution vector within each element is approximated by a tensor product of 1-D Lagrange polynomials ℓ^N of degree N

$$U(\vec{\xi},t) \approx \sum_{i,j,k=0}^{N} \hat{U}_{ijk}(t)\psi_{ijk}^N(\vec{\xi}), \qquad \psi_{ijk}^N(\vec{\xi}) = \ell_i^N(\xi^1)\ell_j^N(\xi^2)\ell_k^N(\xi^3), \tag{15}$$

where $\hat{U}_{ijk}(t)$ are time dependent nodal degrees of freedom and $\ell_i^N(\xi)$ denotes the standard Lagrange polynomial of degree N defined by a nodal set $\{\xi_i\}_{i=0}^N \subset [-1;1]$:

$$\ell_i^N(\xi) = \prod_{j=0;j\neq i}^{N} \frac{\xi-\xi_j}{\xi_i-\xi_j}. \tag{16}$$

A nodal basis offers the advantage of direct knowledge of the interpolant at its nodes, while its counterpart, a modal basis, would require the evaluation of the full basis. In principle, any set of pairwise unique nodes could be chosen to define the interpolation basis in Eq. (16), as long as the resulting interpolation is stable and has a favorable Lebesque constant. The core idea of the DGSEM method is to collocate the interpolation nodes with those that support a quadrature rule of sufficient accuracy. By this choice, the quadrature itself does not require any evaluation of the basis, and - when extending the basis in a tensor-product - becomes a sum of one-dimensional operations in multiple dimensions. Details on this will be demonstrated in Sect. 3.2.

Since the occurring mass matrix is of degree $\sim\xi^{2N}$, the $N+1$ Gauss–Legendre quadrature points $\{\xi_i\}_{i=0}^N$ are chosen as interpolation nodes, as the associated quadrature is exact for this integrand. Another possible choice would be Gauss–Lobatto–Legendre points, leading to a slightly less efficient and accurate scheme due to inexact integration of the mass matrix (Kopriva and Gassner 2010). Now that the approximation of the solution vector U is in place, the discrete transformed flux $\vec{\mathcal{F}}$ can be chosen in a similar manner

$$\mathcal{F}^l(\vec{\xi}) \approx \sum_{i,j,k=0}^{M} \hat{\mathcal{F}}_{ijk}^l \psi_{ijk}^M(\vec{\xi}), \quad l=1,2,3 \tag{17}$$

$$\hat{\mathcal{F}}_{ijk}^l = \mathcal{G}^l(U) - \mathcal{H}^l(U,\vec{\nabla}_x U)\,|_{\vec{\xi}_{ijk}} \tag{18}$$

with $\psi_{ijk}^M(\vec{\xi}) = \ell_i^M(\xi^1)\ell_j^M(\xi^2)\ell_k^M(\xi^3)$. Note that the fluxes are again represented by an interpolation polynomial, but defined on $M+1$ Gauss–Legendre quadrature points,

with $M \geq N$. This implementation allows for a consistent integration of the non-linear fluxes (Kirby and Karniadakis 2003). The choice of M depends on the non-linearity of the flux for under-resolved calculations. For the classical DGSEM, $M = N$ is chosen, which leads to a collocation of solution and fluxes on the same nodes and thus a very efficient implementation.

Now that the domain discretization and the solution and flux approximations have been defined, we can derive the variational formulation of the problem and from it; first the DG formulation, and then the DG discretization scheme. We start by multiplying Eq. (12) by a test function $\phi(\vec{\xi})$ which is taken from the same space as the basis functions. Integrating over the reference element E to leads to the variational formulation in reference space

$$\int_E \left(JU_t + \nabla_\xi \cdot \vec{\mathcal{F}}(U, \nabla_x U) \right) \phi(\vec{\xi})\, d\vec{\xi} = 0. \tag{19}$$

This formulation can be interpreted as an L_2 projection of the residual onto the space of test functions, which enforces orthogonality. Note that so far, no connection to the neighboring elements exist. To remedy this, the second term is rewritten using spatial integration by parts, i.e. the flux divergence is reworked using the product rule of differentiation. Applying the Gauss theorem, the so-called weak formulation of the DG discretization is obtained:

$$\int_E JU_t \phi\, d\vec{\xi} + \oint_{\partial E} \underbrace{\left(\mathcal{G}_n^* - \mathcal{H}_n^*\right)}_{\mathcal{F}_n^*} \phi\, ds - \int_E \vec{\mathcal{F}}(U, \nabla_x U) \cdot \nabla_\xi \phi\, d\vec{\xi} = 0, \tag{20}$$

where $\mathcal{G}_n^*$ denotes the surface normal numerical flux function for the inviscid terms, given by $\mathcal{G}_n^* := \mathcal{G}_n^*(U^+, U^-)$ and superscripts $\pm$ denote the values at the grid cell interface from the neighbor and the local grid cell, respectively. Note that in the volume integral, the flux is no longer required to be differentiable and can be evaluated from information within each grid element, while the new surface integral now contains a numerical flux function to find a unique interface flux from two generally discontinuous left and right states. For the inviscid numerical flux, several well-known flux functions derived for FV formulations are possible, which ensure consistency and uniqueness of the numerical flux. Within the DG community, the most commonly applied flux functions are Godunov's method, the local Lax–Friedrichs or Rusanov flux and Roe's approximate Riemann solver (Toro 1999). The choice of $\mathcal{H}_n^*$ will be discussed in Sect. 3.3.

3.2 The DGSEM Operator

So far, in the derivation of the variational form Eq. (19) and the weak DG formulation Eq. (20), the specific choices made for DGSEM did not come into play. In the

following section, a brief discussion of the DGSEM operator will be given, which is intended to highlight the most important aspects in terms of efficiency. A full, very detailed derivation of the DGSEM operator is given by Hindenlang et al. (2012). As defined in Eqs. (15) and (17), the solution and the flux are represented by tensor products of one-dimensional Lagrange interpolating polynomials, associated with either one-dimensional Legendre–Gauss or Legendre–Gauss–Lobatto quadrature points. The Lagrange property of the basis functions on these nodes makes the evaluation of the basis at these points trivial, as the solution is represented by a nodal interpolation. The evaluation of the inner products is then achieved by the corresponding quadrature rule, which reverts to a sequence of three one-dimensional sums along a reference coordinate line instead of a volume operation including all the element-local solution points. This can be understood as a transfer of the tensor product structure of the basis directly to the operator itself by choosing quadrature and interpolation nodes as described above. This choice reduces the number of operations from $\mathcal{O}(N+1)^6$ for a standard DG formulation to $\mathcal{O}(N+1)^4$ for DGSEM.

We demonstrate this concept of operation reduction by applying the DGSEM formulation to the first volume integral, containing the time derivative of the degrees of freedom, from Eq. (20). First, we insert the ansatz for the solution (Eq. (15)) into the semi-discrete form and choose the test function ϕ from the space of Lagrange polynomials of degree N as ψ^N_{ijk} with associated $N+1$ Legendre–Gauss nodes $\{\xi_i\}_{i=0}^N$

$$\int_E J(\vec{\xi}) U_t \phi \, d\vec{\xi} = \int_E J(\vec{\xi}) \left(\frac{\partial}{\partial t} \sum_{r,s,t=0}^{N} \hat{U}_{rst}(t) \psi^N_{rst}(\vec{\xi}) \right) \psi^N_{ijk} d\vec{\xi} \,. \tag{21}$$

The integral over the reference space is now split into the coordinate directions and then replaced by Legendre–Gauss quadrature with associated weights ω:

$$\begin{aligned}
\int_E J(\vec{\xi}) U_t \phi \, d\vec{\xi} &= \int_{-1}^{1}\int_{-1}^{1}\int_{-1}^{1} J(\vec{\xi}) \left(\frac{\partial}{\partial t} \sum_{r,s,t=0}^{N} \hat{U}_{rst}(t) \psi^N_{rst}(\vec{\xi}) \right) \psi^N_{ijk}(\vec{\xi}) d\vec{\xi}^1 d\vec{\xi}^2 d\vec{\xi}^3 \\
&= \sum_{\alpha,\beta,\gamma=0}^{N} J(\vec{\xi}_{\alpha\beta\gamma}) \left(\frac{\partial}{\partial t} \sum_{r,s,t=0}^{N} \hat{U}_{rst}(t) \underbrace{\ell^N_r(\xi^1_\alpha)}_{=\delta_{r\alpha}} \underbrace{\ell^N_s(\xi^2_\beta)}_{=\delta_{s\beta}} \underbrace{\ell^N_t(\xi^3_\gamma)}_{=\delta_{t\gamma}} \right) \psi^N_{ijk}(\vec{\xi}_{\alpha\beta\gamma}) \omega_\alpha \omega_\beta \omega_\gamma \\
&= \sum_{\alpha,\beta,\gamma=0}^{N} J(\vec{\xi}_{\alpha\beta\gamma}) \frac{\partial}{\partial t} \hat{U}_{\alpha\beta\gamma}(t) \underbrace{\ell^N_i(\xi^1_\alpha)}_{=\delta_{i\alpha}} \underbrace{\ell^N_j(\xi^2_\beta)}_{=\delta_{j\beta}} \underbrace{\ell^N_k(\xi^3_\gamma)}_{=\delta_{k\gamma}} \omega_\alpha \omega_\beta \omega_\gamma \\
&= \underbrace{J(\vec{\xi}_{ijk}) \omega_i \omega_j \omega_k}_{pre-compute} \frac{\partial}{\partial t} \hat{U}_{ijk} \quad \forall \, i,j,k = 0,\ldots,N.
\end{aligned} \tag{22}$$

In Eq. (22), the Kronecker delta functions result from the Lagrange property and thus reduce the associated summation to a single evaluation. Note that the mass matrix is diagonal also for Legendre–Gauss–Lobatto nodes, making the chosen basis both *nodal* as well as *discretely orthogonal*. The Jacobian of the geometry mapping is treated in a collocation way in this approach, i.e. it is not integrated exactly if the mapping is beyond bi-linear. In this case, an additional error akin to mass-lumping for Gauss–Lobatto integration of the mass matrix is introduced.

From an efficiency point of view, Eq. (22) demonstrates how the three-dimensional integrals reduce to point-wise evaluations in DGSEM. For each of the $(N+1)^3$ degrees of freedom $\hat{U}_{ijk}$ per element, just a single multiplication with a pre-computed term is necessary, due to the "folding" of the three-dimensional integral based on the tensor-product structure instead of the evaluation of a three-dimensional integral and inversion of a full mass matrix.

For the surface and flux volume integrals in Eq. (20), a similar reduction in operations can be shown, where the volume integral retains an operation count of $(N+1)$ multiplications per DOF, as the the derivatives of the basis functions do not support the Lagrange property. Further details and a full discretization of the operator can be found in Hindenlang et al. (2012). The semi-discrete form of the DGSEM operator in three dimensions is given as

$$\begin{aligned} -J_{ijk}\left(\hat{U}_{ijk}\right)_t = &\left(\sum_{\alpha=0}^{N} \hat{D}_{i\alpha}\hat{\mathcal{F}}^1_{\alpha jk}\right) + \left([\mathcal{F}^*\hat{s}]^{+\xi^1}_{jk}\,\hat{\ell}_i(+1) + [\mathcal{F}^*\hat{s}]^{-\xi^1}_{jk}\,\hat{\ell}_i(-1)\right) + \\ &\left(\sum_{\beta=0}^{N} \hat{D}_{j\beta}\hat{\mathcal{F}}^2_{i\beta k}\right) + \left([\mathcal{F}^*\hat{s}]^{+\xi^2}_{ik}\,\hat{\ell}_j(+1) + [\mathcal{F}^*\hat{s}]^{-\xi^2}_{ik}\,\hat{\ell}_j(-1)\right) + \\ &\left(\sum_{\gamma=0}^{N} \hat{D}_{k\gamma}\hat{\mathcal{F}}^3_{ij\gamma}\right) + \left([\mathcal{F}^*\hat{s}]^{+\xi^3}_{ij}\,\hat{\ell}_k(+1) + [\mathcal{F}^*\hat{s}]^{-\xi^3}_{ij}\,\hat{\ell}_k(-1)\right)\,, \end{aligned} \tag{23}$$

with the precomputable one-dimensional operators defined as

$$\begin{aligned} D_{ij} &= \left.\frac{d\ell_j(\xi)}{d\xi}\right|_{\xi=\xi_i}, \\ \hat{D}_{ij} &= -D_{ji}\frac{\omega_j}{\omega_i}, \quad i,j = 0,\ldots,N. \end{aligned} \tag{24}$$

The weighted basis functions are given accordingly by

$$\hat{\ell}_i = \frac{\ell_i}{\omega_i}, \quad i = 0,\ldots,N, \tag{25}$$

and $\hat{s}$ is the surface element, relating the physical to the reference surface.

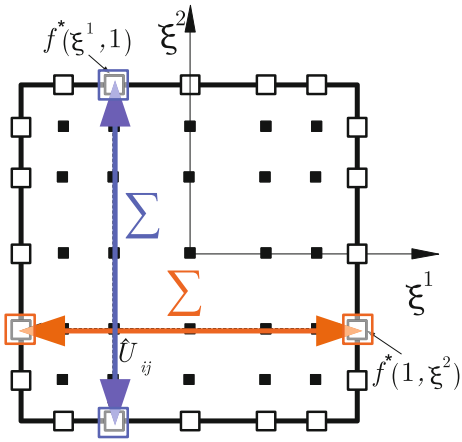


Fig. 6 DGSEM operator structure in 2 dimensions

Equation (23) highlights how the tensor-product basis and collocation of quadrature and interpolation translates to a tensor-product operator and thus becomes computationally highly efficient. The three-dimensional operator essentially collapses to a sequence of three consecutive one-dimensional operators.

Figure 6 visualizes the operator and the involved nodes for two dimensions. The computation of the residual at a given node $\hat{U}_{ij}$ involves essentially three steps: The contribution to the surface integral requires the prolongation of the solution to the element-faces and the subsequent evaluation of the numerical fluxes as a function of the state in the neighboring element. This results in four flux evaluations in 2D. Secondly, the volume contribution is computed by numerical quadrature along two coordinate lines. The third step, the inversion of the mass matrix, is trivially given due to the orthogonality of the basis.

3.3 Approximation of Viscous Fluxes

Returning to Eq. (20), the last missing term to be defined is the numerical approximation for the viscous flux term $\mathcal{H}_n^*$. This term introduces a dependence on the gradient of the solution. The treatment of the gradient terms in the context of DG approximations was first tackled by Bassi and Rebay (1997), who introduced a mixed finite element approximation, in which the gradients are approximated in the same discontinuous polynomial space as the solution. They also showed that a local evaluation of the gradient leads to instabilities, and that some form of "lifted" gradient, containing information from both adjacent elements, is needed.

To derive the mixed formulation, the system of governing equations is rewritten as a corresponding system of first order equations with an auxiliary variable $\vec{S}$ as an approximation of the lifted gradients

$$\begin{aligned} \vec{S} - \nabla_x U &= 0, \\ U_t + \nabla_x \cdot \vec{F}(U, \vec{S}) &= 0. \end{aligned} \tag{26}$$

Applying the discretization steps outlined above to derive a weak DG discretization of the auxiliary equation leads to

$$\begin{aligned} c = 1, \ldots, 5: \quad & \int_E J \vec{s}_c \phi \, d\vec{\xi} + \oint_{\partial E} \vec{u}^*_{c,n} \phi \, ds - \int_E u_c \cdot \nabla_\xi \phi \, d\vec{\xi} = 0, \\ & \int_E J U_t \phi \, d\vec{\xi} + \oint_{\partial E} \left(\mathcal{G}^*_n - \mathcal{H}^*_n \right) \phi \, ds - \int_E \vec{\mathcal{F}}(U, \vec{S}) \cdot \nabla_\xi \phi \, d\vec{\xi} = 0, \end{aligned} \tag{27}$$

with the component u_c of the state vector U and its lifting operator $\vec{s}_c$. The numerical flux of the auxiliary equation is $\vec{u}^*_{c,n}$, and $\mathcal{H}^*_n = \mathcal{H}^*_n(U^+, U^-, \vec{S}^+, \vec{S}^-)$ denotes the numerical flux function for the viscous terms. Following Bassi and Rebay (1997), we choose

$$c = 1, \ldots, 5: \quad \mathfrak{u}^*_{c,n} = \left(\alpha_{visc} \, u^+_c + (1 - \alpha_{visc}) \, u^-_c \right) \vec{n}, \tag{28}$$

$$\mathcal{H}^*_n = \left(\alpha_{visc} \, \mathcal{H}_n(U^+, \vec{S}^+) + (1 - \alpha_{visc}) \, \mathcal{H}_n(U^-, \vec{S}^-) \right), \tag{29}$$

with $\vec{n}$ denoting the outward pointing surface normal. For a parameter of $\alpha_{visc} = \frac{1}{2}$, this treatment of the viscous fluxes is usually labeled BR1 (first method of Bassi and Rebay 1997).

3.4 Boundary Conditions

Since in DG methods, the coupling between the elements is achieved weakly or indirectly through the numerical flux as a function of the adjacent states, it is natural to extend this approach to the boundary conditions as well. The rationale for this approach is to ensure consistency in the approximation of the internal faces fluxes and the boundary conditions, i.e. to use the same discretization operators for both and thus avoid stability issues (Bazilevs and Hughes 2007). This approach is also applicable to Dirichlet type boundaries, where instead of prescribing a state U directly at the boundary, an appropriate right hand side state U^+ (akin to a ghost cell state) is prescribed. Together with its adjacent neighbor state from within the domain, it is then used to compute the resulting advection boundary flux through the appropriate Riemann solver. The gradients for the diffusive fluxes are chosen according to the specific type of boundary condition.

Collis (2002) investigated the effect of weakly versus strongly imposed Dirichlet conditions for the case of an under-resolved one-dimensional stationary boundary layer problem and turbulent channel flows. He found that the L_2 error norm is greatly

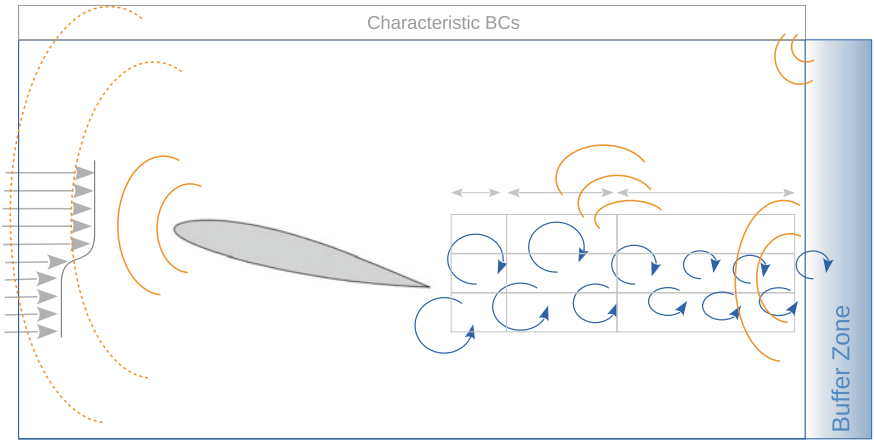


Fig. 7 Boundary conditions and acoustic disturbance sources

reduced when the boundary conditions are enforced in this weak manner and oscillations near the boundary are avoided. We follow this approach, and we enforce the boundary conditions for Eq. (27) weakly through the prescription of the right hand side state U^+ of the boundary fluxes $\vec{\mathfrak{u}}^*_{c,n}$, $\mathcal{G}^*_n$ and $\mathcal{H}^*_n$.

Challenges in Aeroacoustics Due to the scale separation between hydrodynamics and acoustics, stable and accurate boundary conditions pose a considerable challenge. In particular, outflow and far field conditions are difficult to handle in subsonic flows, since they are usually artificial boundaries and thus the correct outer state is not known. Especially at the outflow boundary, where large scale non-linear hydrodynamic structures need to exit the domain, a slight error in the boundary condition will act as any gradient in the Lighthill tensor and produce noise radiating from the boundaries into the domain. Figure 7 highlights some of the challenges of applying boundary conditions and sources that may pollute the acoustic field.

Another issue that is not directly related to the boundary condition treatment is the generation of sound waves at gradients of the local resolution, e.g. at stretched or skewed grid cells where waves become more poorly resolved and their non-resolvable energy is radiated again as acoustics. This becomes particularly troublesome for high order discretization with low numerical dissipation.

Across the inflow boundary, the incoming flow state and possible noise disturbances are described. Upstream propagating waves must be able to leave the domain without reflections. Recognizing that these waves are typically of low amplitude, boundary condition types based on linearization and characteristic decomposition work well. This also holds for the parts of the boundary that are approximately parallel to the flow, as long as no large amplitude noise occurs. At the outflow boundary, approaches based on linearization are generally not successful when high amplitude disturbances like turbulent structures encounter the boundary. Without special treatment, the outflow boundary can act as a dominating artificial source and pollute the

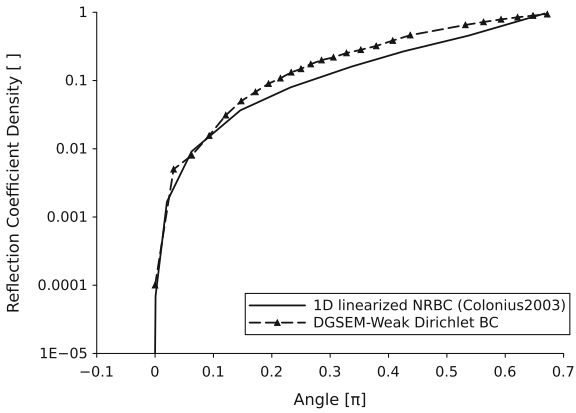


Fig. 8 Reflection coefficient of different boundary conditions

whole acoustic field. No general theoretical method exists to construct non-reflecting boundary conditions in this situation. Instead, an artificial absorbing layer is the most commonly used approach. This layer is placed upstream of the boundary itself and modifies the incoming flow and acoustic field while it is passing through it. The amplitudes of the disturbances are damped towards a "quiet" base flow, which then exits through a linearized boundary condition. While this approach can be very effective and computationally efficient, it introduces a number of user-selectable parameters such as layer width or ramping function and itself can also become reflective. In addition, careful blending of the buffer region with other adjacent boundaries must be implemented to avoid generation of disturbances through a mismatch. Colonius (2004) gives a good overview of possible implementations for ad-hoc solutions, such as sponge zones, perfectly matched layers, fringe and grid stretching.

Boundary Conditions for DG In the following section, we will briefly discuss how the typical boundary conditions discussed above can be treated in the DG context.

Far Field Boundaries As discussed above, the boundary conditions in DG are prescribed weakly through a numerical flux. The choice of this flux function can be adapted to the problem at hand; for the acoustic far field, those that are based on a wave decomposition are a natural choice. The state outside of the domain is fixed to the free-stream state at infinity, and the resulting boundary flux is computed with the adjacent inner state. It has been shown in Flad et al. (2014) that employing Roe's Riemann solver flux function mimics classical one-dimensional linearized characteristic non-reflecting boundary conditions and effectively prevents reflections of acoustic disturbances.

Figure 8 shows the reflection behavior of a planar acoustic wave transported with background velocity $u_0 = 0.5$ and speed of sound $c = 1$, crossing the boundary under different incident angles. It is compared to a 1D linearized local boundary condition proposed by Colonius (2004). The reflection coefficient $max(|\rho_{refl}|)/max(|\rho|)$ vanishes for small angles and drops below 5% for angles smaller than $\approx 25°$.

Outflow Boundaries A simple and robust variant of the absorbing layer discussed above is the sponge zone concept, in which a retarding volume source term is included in the spatial operator

$$U_t = R(U) - d\sigma(\vec{x})(U - U_B), \tag{30}$$

where $R(U)$ represents the discretized Navier–Stokes operator, d controls the magnitude of the source term and $\sigma(\vec{x})$ denotes a ramping function. This ramping function is intended to prevent reflections at the domain - sponge interface and smoothly increases the source term strength towards the domain edges. U_B identifies an acoustically quiet base flow towards which the solution is forced; clearly, if $U_B = U$, the source term vanishes. Suitable choices for U_B are a constant free-stream state or a time-averaged solution from a prior simulation. A flexible and general method to determine a suitable base flow is to generate it from a moving time-average of the solution. This time-average is computed by an exponential temporal filter, which can be written in a simple differential form as

$$\bar{U}_t(t, \Delta) = \frac{U(t) - \bar{U}(t, \Delta)}{\Delta}, \tag{31}$$

This expression only requires storing of one previous base flow states. It is integrated in time alongside the spatial DG operator and thus yields the base flow $U_B = \bar{U}(t, \Delta)$ in every time step. This filter idea has been adopted from the temporally filtered LES by Pruett et al. (2003). The filter width Δ should be set to cover the largest time scales of the flow.

Following Flad et al. (2014) to demonstrate the effectiveness of this sponge method, a 2D isentropic Euler vortex with an initial maximum density perturbation of 13.3% is transported at $Ma = 0.5$ along the x-direction. It is computed without any damping zone and with the adapting sponge approach discussed above. The sponge layer uses a ramping function of width Δx_{SP} which starts at x_0 and uses a polynomial blending given by

$$\sigma(x^*) = 6x^{*5} - 15x^{*4} + 10x^{*3}, \tag{32}$$

and with

$$x^* = \frac{x - x_0}{\Delta x_{SP}} \tag{33}$$

being the local sponge coordinate. The parameters for the sponge zone are $x_0 = 10$, $\Delta x_{SP} = 10$, $d = 0.1$, $\sigma(x^*)$ equals 0.1 for $20 \leq x \leq 25$ and $\Delta = 20$. Figure 9 gives a qualitative impression of the reflection entering the computational domain of interest $x < 10$ without and with a sponge zone, showing a significant reduction of reflected acoustics for the latter case.

As shown by Akervik et al. (2006), using a moving temporal average as a base state has the additional advantage of not altering the steady state solution. This implies that the sponge zone can be initiated very closely to the region of interest, which reduces computational costs compared to other zonal concepts. In addition, this reduces the

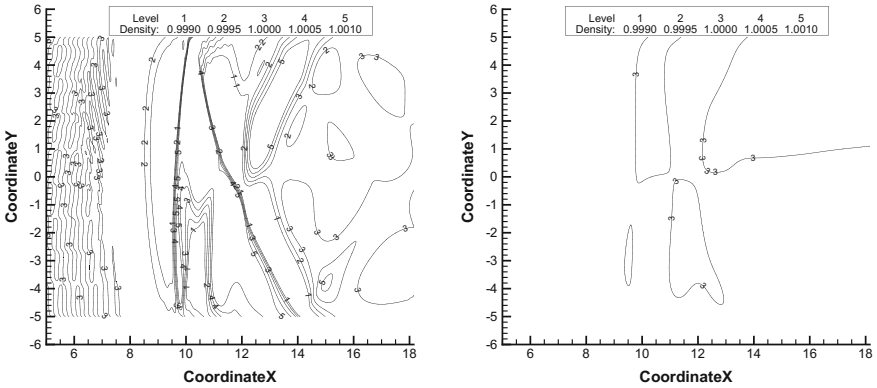


Fig. 9 Density contours at $t = 74$. *Left* without zonal BC, *Right* with adaptive sponge zone

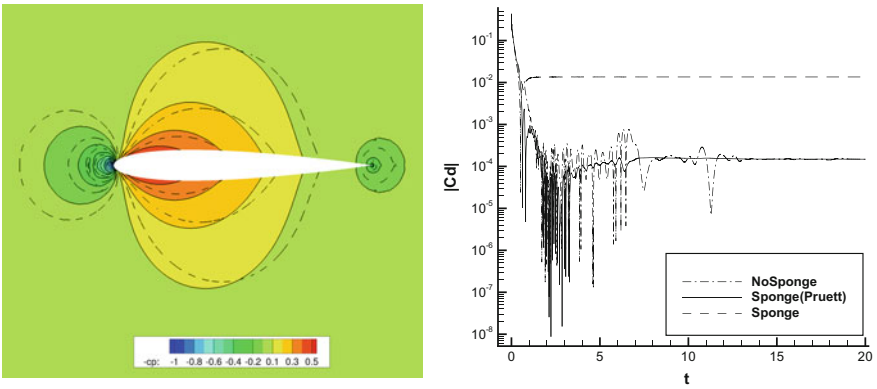


Fig. 10 *Left*: c_p contours: no sponge (*colors*), Pruett sponge (*solid lines*), sponge (*dashed lines*), *Right*: c_d convergence

sensitivity with regards to the sponge zone parameters and thus removes a source of computational uncertainty. To demonstrate this feature of the adjusting base flow, Fig. 10 (left) shows the results of a 2D Euler flow simulation at $Ma = 0.4$ around a NACA 0012 airfoil. Three simulation results are compared: (1) One without any sponge zone, one with a constant sponge (2) and one with the adjusting one (3), both applied in the entire field. The damping parameter is set to $d = 0.1$ and the filter width is 0.5 convective times (c/u_∞). Solution (1) and (3) show an identical flow field, while the classical sponge (2) clearly influences the steady state solution due to the base flow inconsistency. The right pane of Fig. 10 shows the convergence of the drag coefficient, which differs from the unfiltered results for case (2). Thus, the adjusting sponge zone can be implemented efficiently without a large memory requirement, it retains the steady state solution as it filters in time and it can be used to prevent reflections.

The full framework described in this section is available as the open-source code package FLEXI[1] under GPL 3.0.

4 Applications

In this section, we describe some CAA simulations with the DGSEM method presented in Sect. 3. We start by a brief presentation of a LEE sound scattering simulation to highlight the influence of the numerical scheme in terms of the n_{ppw} criterion in Sect. 4.1. In Sect. 4.2, tonal noise generation at an airfoil is computed and compared to well-established results to validate the established framework. In Sect. 4.3, we present the simulation of a feedback mechanism in a complex automotive test case.

4.1 Linearized Euler Equations

While the framework presented in Sect. 3 is mainly intended for direct methods, implementing hyperbolic/parabolic systems of equations beyond the compressible Navier–Stokes equations is straight forward. For this investigation, the Linearized Euler Equations (LEE) have been implemented:

$$\frac{\partial \rho}{\partial t} + (\mathbf{v}_0 \cdot \nabla_x)\rho + \rho_0 \nabla_x \cdot \mathbf{v} + (\mathbf{v} \cdot \nabla_x)\rho_0 + \nabla_x \cdot \mathbf{v}_0 \rho = 0 \tag{34}$$

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v}_0 \cdot \nabla_x)\mathbf{v} + \frac{1}{\rho_0}\nabla_x p + (\mathbf{v} \cdot \nabla_x)\mathbf{v}_0 + \frac{1}{\rho_0}(\mathbf{v}_0 \cdot \nabla_x)\mathbf{v}_0 \rho = 0 \tag{35}$$

$$\frac{\partial p}{\partial t} + (\mathbf{v}_0 \cdot \nabla_x)p + \gamma p_0 \nabla_x \cdot \mathbf{v} + (\mathbf{v} \cdot \nabla_x)p_0 + \gamma(\nabla_x \cdot \mathbf{v}_0)p = 0 \tag{36}$$

For a constant base state $U_0 = (\rho_0, u_0, v_0, w_0, p_0)$, they can written in conservative form as

$$U_t + \mathbf{A}\, U_x + \mathbf{B}\, U_y + \mathbf{C}\, U_z = S, \tag{37}$$

with the acoustic source term S and the matrices **A, B, C** depending on U_0 only. Following the test case description from the Second Computational Aeroacoustics Workshop on Benchmark Problems (Tam and Hardin 1997), the scattering of a point source on a cylindrical object of diameter $D = 1$ is investigated. The source term acts through periodical pressure and density disturbances and is given by

[1] www.flexi-project.org.

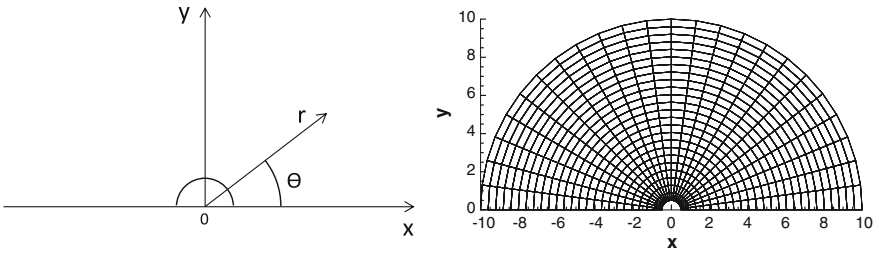


Fig. 11 Grid and coordinate system for LEE cylinder scattering test case

$$S = \begin{pmatrix} \frac{\gamma}{c_0^2} S(x,y,t) \\ 0 \\ 0 \\ 0 \\ S(x,y,t) \end{pmatrix}, \; S(x,y,t) = \exp\left[-ln(2)\frac{(x-x_c)^2 + (y-y_c)^2}{b^2}\right] \sin(\omega t), \tag{38}$$

with $b = 0.2$, $\omega = 8\pi$, $\gamma = 1.4$ and the source coordinates $x_c = 4$ and $y_c = 0$. Initially, $U(t = 0) = 0$ and the background state is given by

$$\rho_0 = 1, \mathbf{v}_0 = \mathbf{0}, p_0 = 0.714285714. \tag{39}$$

The computational domain is a 2D half cylinder of radius $r = 10$, discretized by a structured grid with refinement towards the geometry. Symmetry boundary conditions are enforced on the lower boundary, while Dirichlet boundaries with vanishing fluctuations are enforced on the outer surface. Figure 11 depicts the geometry and grid as well as the coordinate system used. A number of computations have been conducted on different hierarchical grids and varying polynomial degree N. As a numerical flux function, a standard characteristic flux vector splitting was used. Figure 12 shows the instantaneous pressure and velocity fluctuations at $t = 100$.

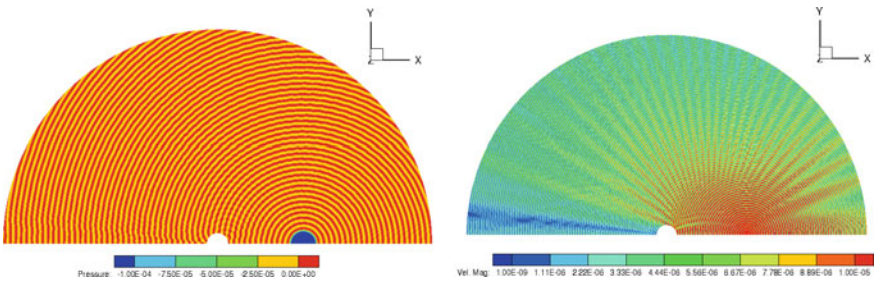


Fig. 12 Instantaneous pressure and velocity fluctuations at $t = 100$

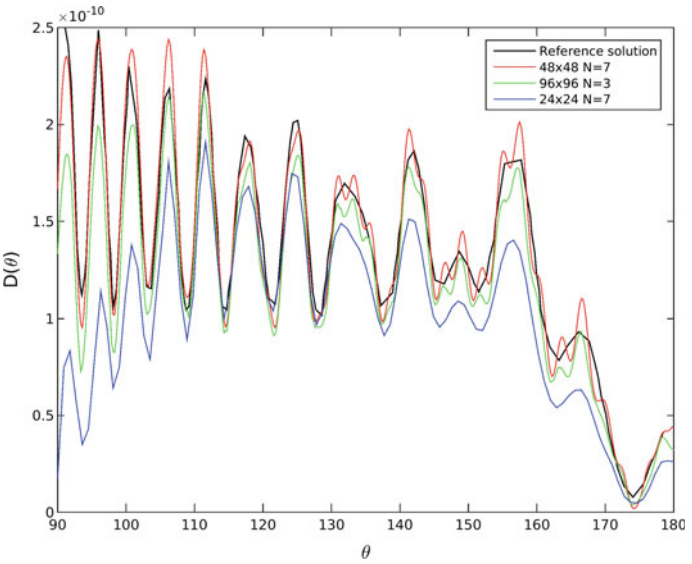


Fig. 13 Comparison of computed directivity with reference solution from Tam and Hardin (1997)

For a quantitative comparison, the directivity D as a function of the radius r can be computed from

$$D(\theta, r) = r\overline{p(\theta)^2}, \tag{40}$$

where the bar denotes the time-averaging. The averaging takes place between $t = 45$ and $t = 100$ once the initial disturbances have left the domain. Figure 13 compares the simulation results with analytical reference data. The simulations were conducted on three hierarchically refined structured grids with 24×24, 48×48 and 96×96 elements in the $x - y$-plane. The degree of the polynomial approximation was chosen to be $N = 3$ and $N = 7$. From the results in Fig. 13, two trends can be observed. Firstly, the solution improves towards the reference solution when the grid size is halved while keeping N constant. Secondly, the importance of a low n_{ppw} criterion is demonstrated here, as two simulations with the same overall number of DOF (96 elements, $N = 3$ and 48 elements, $N = 7$) differ significantly in accuracy. For this nominal resolution (taken along a 1D line at the lower boundary), the number of DOF per acoustic wavelength is $\approx$6. From the discussion in Sect. 2.2, this is less than optimal for accurate wave representation for an $N = 3$ approximation, but sufficient for $N = 7$. Accordingly, the $N = 7$ solution is in better agreement with the analytical reference. Figure 14 supports these observations. For the same total number of DOF, the high order solution (left) retains the acoustic waves up to the boundary, while for the low order solution (right), only the waves in close proximity to the source are kept intact.

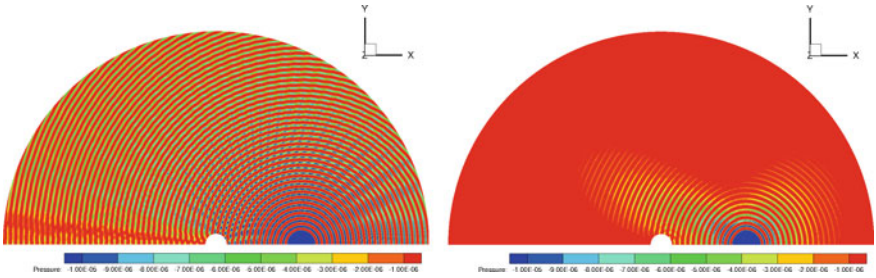


Fig. 14 Instantaneous pressure fluctuations at $t = 100$ with 384^2 DOF. *Left*: 48×48 elements, $N = 7$; *Right*: 192×192 elements, $N = 1$

4.2 NACA 0012 Tonal Noise

The DGSEM framework described in Sect. 3 has been applied to a number of turbulent and transitional test cases (Fechter et al. 2012; Flad et al. 2014; Beck et al. 2016, 2014) covering laminar separation, transition and turbulent reattachment in an LES setting. In the following, we will discuss the simulation of the flow around a NACA 0012 airfoil which has been shown to support the establishment of an acoustic feedback loop (Paterson et al. 1973; Arbey and Bataille 1983; Nash et al. 1999; Desquesnes et al. 2007; Jones and Sandberg 2011; Plogmann et al. 2013). We follow the 2D DNS of Jones and Sandberg (2011) and conduct a well-resolved simulation at $Ma = 0.4$ and $Re_C = 100{,}000$ based on the chord C at an angle of attack of $\alpha = 0°$.

The 2D domain is discretized in a C-type topology. The upstream radius is $r = 7C$, and extends to $9C$ downstream. The domain is divided into 40,934 unstructured elements, each supporting a polynomial of degree $N = 5$ per direction. This results in about 1.5 million degrees of freedom. The boundary geometry is represented by a polynomial of degree $N_{\text{geo}} = 4$ per direction. This ensures proper representation of the airfoil curvature. Details on the near-wall resolution of the current and the reference simulation from Jones and Sandberg (2011) are listed in Table 2. The far-field boundary conditions are enforced weakly, with a Roe Riemann flux function to enable the exiting of low amplitude waves as discussed in Sect. 3.4. In addition, a circular moving-average sponge zone is arranged around the trailing edge, with its

Table 2 Wall-tangential and wall-normal grid spacing Δx and Δy at the leading edge (LE) and trailing edge (TE) for the current simulation of the NACA 0012 case and reference Jones and Sandberg (2011). $\Delta x = \Delta x_{\text{Elem}}/(N+1)$, $\Delta y = \Delta y_{Elem}/(N+1)$

	LE current	LE ref.	TE current	TE ref.
$\Delta x/C$	$4.2 \cdot 10^{-4}$	$6.1 \cdot 10^{-4}$	$5.3 \cdot 10^{-4}$	$4.0 \cdot 10^{-4}$
$\Delta y/C$	$2.3 \cdot 10^{-4}$	$3.5 \cdot 10^{-4}$	$2.3 \cdot 10^{-4}$	$4.0 \cdot 10^{-4}$

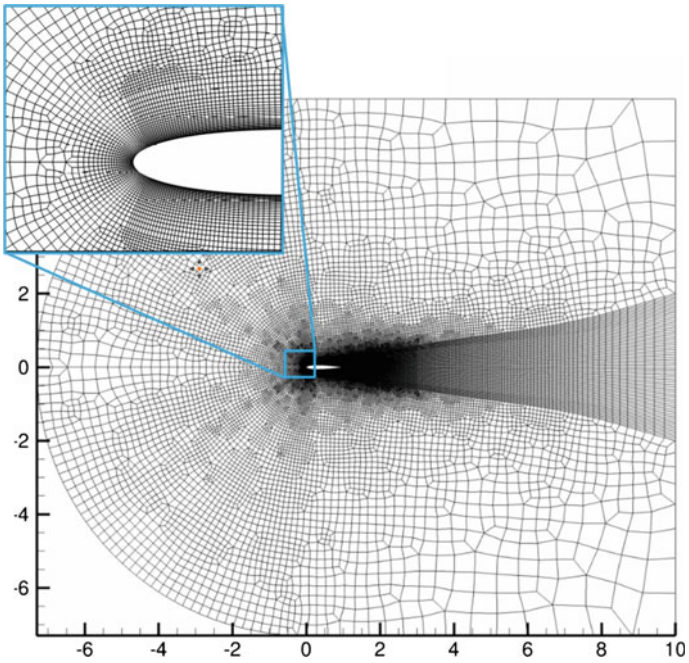


Fig. 15 Domain and grid for the NACA 0012 simulation

source term strength $d\sigma(\vec{x})$ ramped from 0 to 0.5 in the range $r/C = 2$ to 6, while the temporal filter width is set to $\Delta = 2C/u_\infty$. Figure 15 shows the domain and the grid. The simulation was conducted on the CRAY XC40 Hornet cluster using 720 cores, which resulted in a load of ≈2000 DOF/core, which is near the optimum of the framework. The resulting computational wall time amounted to 3 min per convective time unit $T^* = C/u_\infty$ at a time step of $\Delta t/T^* = 3.1 \cdot 10^{-5}$.

The general flow features are illustrated by instantaneous vorticity contours in Fig. 16 (left). The boundary layer separates on both sides of the airfoil, which leads to a roll-up of vortices slightly upstream of the trailing edge. Figure 16 (right) shows the associated acoustic radiation by means of volume dilatation ($\vec{\nabla}_x \cdot \vec{v}$) contours. The typical dipole character of trailing edge noise can be easily recognized.

The acoustic signal at an observer position of $0.5C$ above the airfoil can be compared to the reference in Fig. 17 by means of the PSD of pressure. The PSD is approximated by averaging over 5 blocks with 50% overlap and a Hanning window over a total of $36T^*$. The main tonal frequency and the overall shape of the decaying broadband noise are in excellent agreement. Deviations are found in the missing side peak at $fT^* \approx 3.3$ and additional lower side peaks at $fT^* \approx 2$ and 2.9 yielded by the present simulation, which do not appear in the reference.

In order to determine whether the underlying feedback mechanism is present and detected by the numerical simulation, a global stability analysis is conducted. More

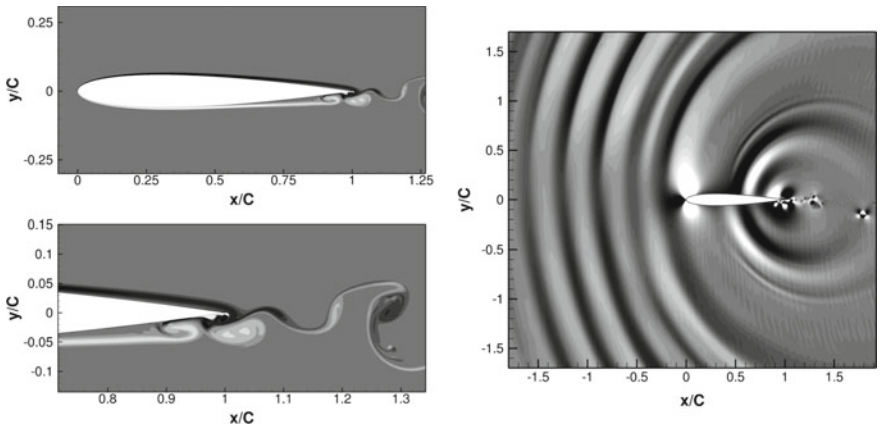


Fig. 16 *Left*: instantaneous vorticity contours over the range $\Omega_z = \pm 100 u_\infty / C$, *Right*: volume dilatation contours in the range $\vec{\nabla}_x \cdot \vec{v} = \pm 0.1 u_\infty / C$

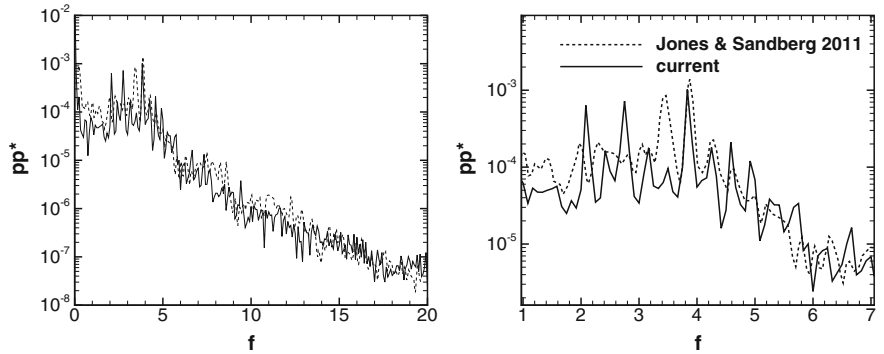


Fig. 17 PSD of pressure at $\vec{x}/C = (0.5, 0.5)$, where the origin is placed at the leading edge of the airfoil

details on this method can be found in Frank and Munz (2016). The basic concept of this type of analysis is to consider the temporal evolution of small disturbances on a frozen base flow. To this end, the solution U is rewritten as a Reynolds decomposition of the form $U = U_0 + U'$, where $(U_0)_t \neq 0$ for a general base flow. Introducing this ansatz into the evolution equation leads to an expression for the dynamics of small perturbations to arbitrary base flows:

$$U'_t = R(U_0 + U') - R(U_0). \tag{41}$$

This simple perturbation formulation is suitable for any non-linear solver, as it just requires the subtraction of the operator evaluated at the base state at every instance. A Taylor series expansion of the full evolution equation about U_0 shows that Eq. (41) approximates a linearization for small U'.

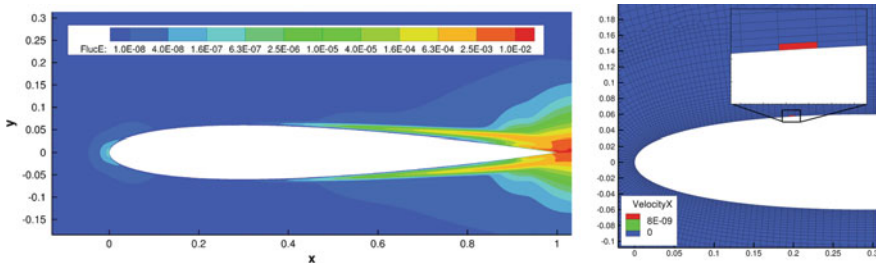


Fig. 18 *Left*: RMS fluctuations of velocities $\sqrt{\overline{u'u'} + \overline{v'v'}}$ for NACA 0012. *Right*: location of perturbation

For the present analysis, we chose the time-averaged base flow as U_0. It was then initially perturbed by a cell-constant value of $U'(t = 0)/U_\infty = 10^{-8}$ and left to evolve according to Eq. (41).

Figures 19 and 20 visualizes the feedback loop as an interplay of hydrodynamic instabilities (visualized by the vorticity in the left column) and the acoustic field (shown is the dilatation in the right column). Starting from the top, the perturbation was introduced at time $t = 0$ at the location shown in Fig. 18 (right). The perturbation is convected along the airfoil and grows in amplitude in the separated shear layer (t_1 in Fig. 19). As it passes the trailing edge, large scale acoustic radiation is generated and propagates (also) upstream (t_2). At t_3, the energetic part of the wave package has left the trailing edge, and the associated acoustic radiation subsides, leading to a visually "quiet" state again (t_4 in Fig. 20). Some time later, although no further perturbation has been introduced externally, a new energetic wave package appears (t_5), which again generates acoustics upon shedding (t_6), thereby closing the loop.

This simulation of the feedback loop and comparison with published results serves as a validation case for our framework. The accurate prediction of the acoustic signal and the establishment of the feedback loop are only possible if the precise hydrodynamic and acoustic processes are captured by the simulation. The close agreement of our simulation with the acoustic results of the reference demonstrate the suitability of our high order code framework for aeroacoustic feedback effects. In the following, it will be applied to a more complex case of acoustic feedback at an automotive side mirror.

4.3 Acoustic Feedback Mechanism at a Side Mirror

Feedback as a source of tonal noise In this section, we will present the application of the framework to a complex acoustic problem in an industrial setting, namely the tonal noise generated by an acoustic feedback loop on a car side mirror alongside experimental data. The numerical results presented in here are based on the work by Frank (2016), while the joint experimental analysis was conducted by Werner (2017).

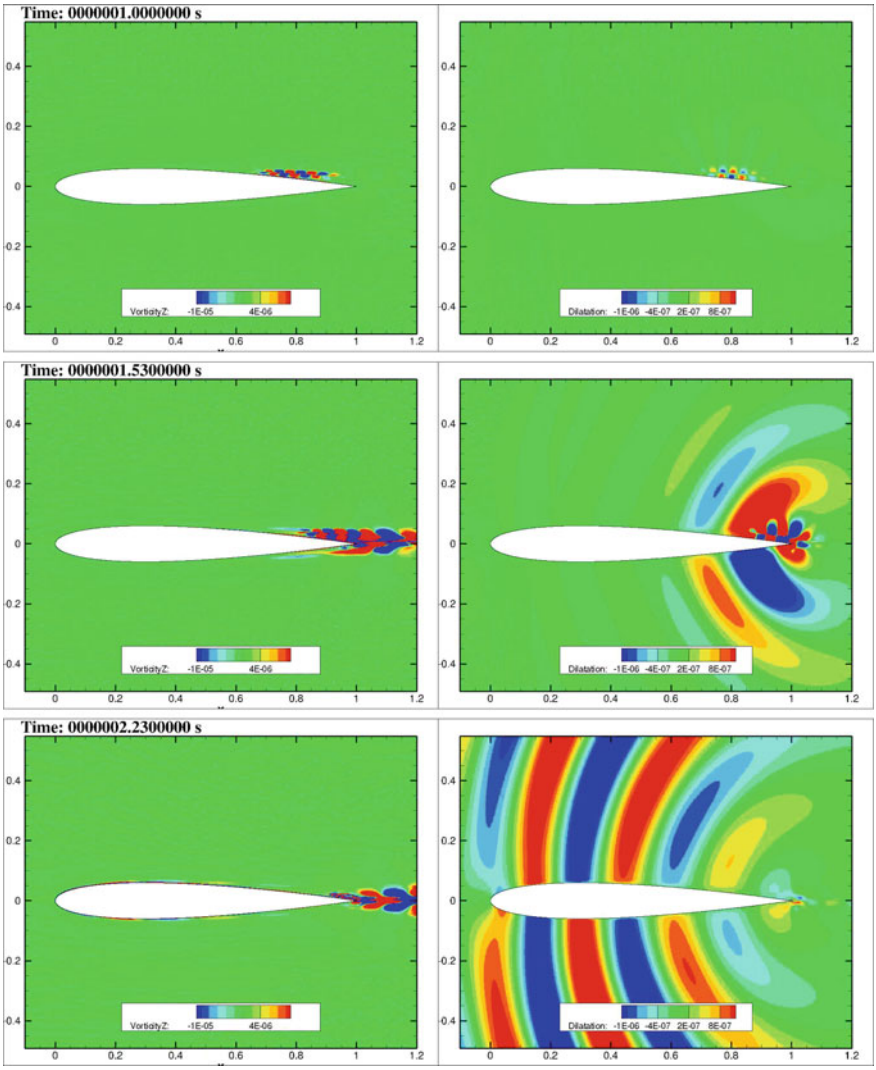


Fig. 19 Part A: temporal evolution of z-vorticity and dilatation rate for NACA 0012 case, each row corresponds to an instance in time t_i, $i = 1, \ldots, 3$. (See also Fig. 20)

Aerodynamically, typical mirror shapes can be classified as bluff-body configurations, with the associated flow phenomena. As observed experimentally, the mirror is known to be a source of tonal noise. The associated narrowband amplitude peaks in the acoustic spectrum are typically perceived as disturbing whistling sounds. The main method of research into their origins for the mirror configuration remains experimental, and is limited to simplified mirror geometries for numerical studies. The only simulation of a realistic mirror known to the authors was reported by

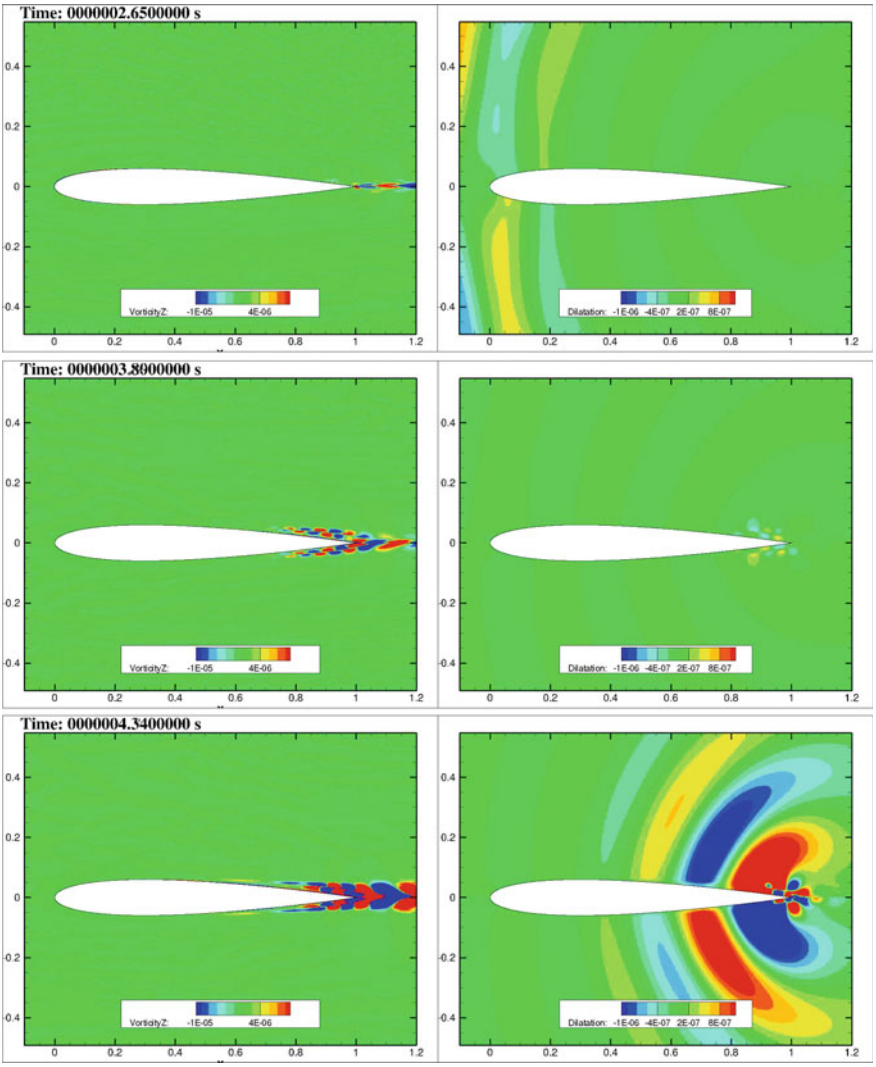


Fig. 20 Part B: temporal evolution of z-vorticity and dilatation rate for NACA 0012 case, each row corresponds to an instance in time t_i, $i = 4, \ldots, 6$. (See also Fig. 19)

Khalighi et al. (2010). For more generic geometries like airfoils however, a number of numerical simulations of self-noise exist, e.g. Jones and Sandberg (2010), Desquesnes et al. (2007), Chong and Joseph (2012), see also the computation presented in Sect. 4.2. Lounsberry suggested that a similar feedback mechanism as the one found on airfoils was responsible for the noise generation along car mirrors, noting the shared occurrence of attached laminar or transitional boundary layers up until close to the trailing edge (Lounsberry et al. 2007).

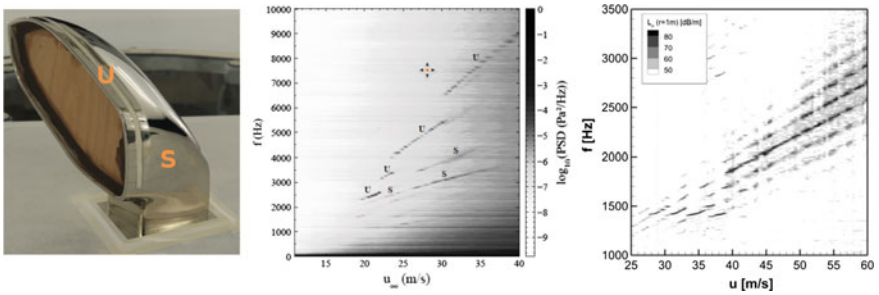


Fig. 21 *Left*: model mirror on wind tunnel floor from Werner et al. (2017a), *Middle* acoustic measurements on mirror, from Werner et al. (2017b), *Right*: acoustic measurements on NACA 0012 airfoil, from Plogmann et al. (2013)

While tonal noise at airfoils has been observed both experimentally e.g. Arbey and Bataille (1983), Plogmann et al. (2013), Paterson et al. (1973) as well as numerically, different theories about the exact mechanism exist. Paterson et al. (1973) attributed the noise to the bluff-body vortex shedding with a distinct Strouhal frequency at the trailing edge. However, this explanation did not account for the ladder-type structure observed when plotting the tonal frequency over the freestream velocity, i.e. the distinct jumps in frequency, see Fig. 21. Several other models have been proposed, which are based on the concept that for a self-sustaining feedback loop, the phase difference over one cycle should vanish, a condition that only discrete frequencies can fulfill (Tam 1974; Kingan and Pearse 2009; Arbey and Bataille 1983). Using receptivity strips at different locations in the laminar boundary layer along a NACA 0012 airfoil, Plogmann et al. were able to trigger receptivity experimentally, which resulted in a change of the tonal frequency according to the phase criterion, strongly supporting the notion of acoustic feedback as an explanation for the frequency selection.

Figure 21 (left) shows the mirror model on the floor of the Laminar Wind Tunnel at the Institute of Aerodynamics and Gasdynamics (IAG). The measured frequency spectra as a function of freestream velocity u_∞ are shown alongside (middle plot). The ladder structure is visible both for the side and upper surfaces. Through boundary layer tripping, the regions of tonal noise generation could be established. For comparison, the right plot depicts similar measurements for a NACA 0012 airfoil.

In Fig. 22, the building blocks of the feedback loop are shown: A laminar boundary layer along a convex geometry separates close to the trailing edge due to the adverse

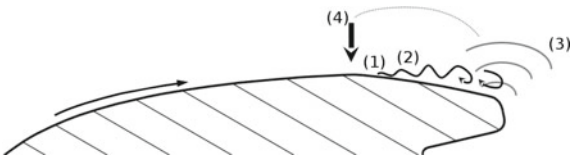


Fig. 22 Conceptual model for the feedback loop

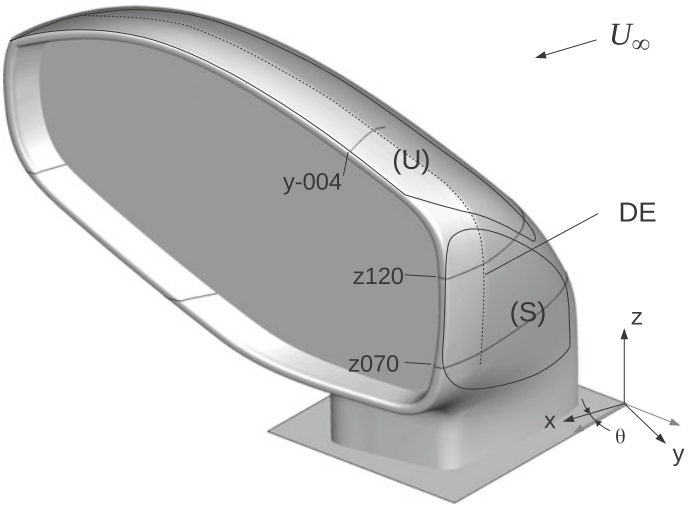


Fig. 23 Mirror geometry. Marked areas are *S* Side surface, *U* Upper surface, *DE* Design edge

pressure gradient (1). In the resulting detached shear layer, convective instabilities are amplified and start the roll-up into coherent vortices (2). When passing the trailing edge, these structures generate sound waves through scattering (3). Pressure waves run upstream through the boundary layer and reinforce the boundary layer instability due to receptivity, thereby closing the loop (4). This is essentially the same mechanism as presented in Sect. 4.2.

Numerical Model As confirmed by the experimental investigation of Werner et al., the non-generic early-development-stage side-view mirror depicted in Fig. 21 develops a distinct whistling sound at normal cruise speeds, provided that the inflow turbulence level is kept low (Werner 2017). Figure 23 shows the computational model of this mirror geometry and the coordinate system. To enable direct comparison with the experimental data, an isolated mirror was considered. The length scale $L = 0.1\,\mathrm{m}$ corresponds to the lateral length of the side surface. The free-stream velocity was set to 100 km/h, and a yaw angle of $\theta = -20°$ was chosen. This angle resulted from a preliminary investigation, in which it was found that this yaw angle resulted in a comparable pressure distribution on the mirror side surface, when compared to a full configuration with the mirror mounted on the car chassis. As the mirror side and upper surface are outside the wind tunnel boundary layer (Frank 2016), no influence of the wind tunnel boundary layer on the tonal noise generation is expected. Therefore, symmetry boundary condition are applied on the wind tunnel floor, while the free-stream boundary conditions are chosen as weakly enforced Dirichlet conditions, see Sect. 3.4. On the mirror geometry, isothermal wall boundary conditions are applied. A temporally adapting sponge zone is added upstream of the outflow boundary. The associated source term is ramped parallel to the free-stream velocity vector beginning at approximately $2L$ downstream of the average trailing edge of the mirror. Based

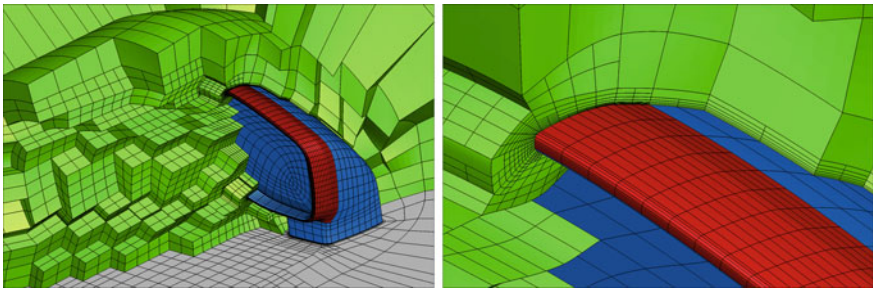


Fig. 24 Cut view of the computational mesh close to the mirror, showing the non-conforming interfaces

on the time scale $T = L/u_\infty$, the damping parameter and the temporal filter width are set to $d = 0.8/T$ and $\Delta = 4T$, respectively.

The computational mesh is created in two-step process. First, a coarse, block-structured grid made of hexahedral cells is generated with a commercial grid generator. Afterwards, this mesh is refined in user-specified regions by isotropic cell splitting, namely in the boundary layer around the mirror and in the wake region up to the outflow boundary. This introduces non-conforming cell interfaces. A second refinement level is introduced on the trailing edge area, in which the feedback process outlined in Fig. 22 is known to occur from the experiments. This area is marked in red in Fig. 24. The refinement is managed by the software library p4est (Burstedde et al. 2011), the resulting mesh consisted of 32,800 elements. The curved surface of the mirror is realized via an agglomeration approach (Hindenlang et al. 2015). The mapping from reference to physical space is constructed from a super-sampled version of the unrefined base grid. For the refined cells, the mapping can simply be evaluated in the respective subset of the lower level parameter space. This way, we ensure free stream preservation and conservation also at the non-conforming interfaces.

Simulation Results

Spatial resolution To assess the influence of the spatial resolution on the results, a p-refinement is conducted by increasing the local polynomial degree. This not only increases the number of degrees of freedom, but also shifts the n_{ppw} factor due to the increase in the approximation order. Two resolutions are considered: Case $N4M6$ denotes an approximation of degree $N = 4$, with an evaluation of the non-linear inner products with an approximation of degree $M = 6$. Analogously, case $N7M10$ denotes an approximation of degree $N = 7$. Details on the de-aliasing approach can be found in Beck (2015). The resulting mesh parameters are listed in Table 3.

Table 3 Computational mesh and resolution details

Case	DOF	Δt [s]	Δy [mm]	Δy^+	Δx [mm]	Δx^+
N4M6	$4.1 \cdot 10^6$	$3.9 \cdot 10^{-8}$	0.026	2.5	0.6	80
N7M10	$16.8 \cdot 10^6$	$1.8 \cdot 10^{-8}$	0.016	1.5	0.38	40

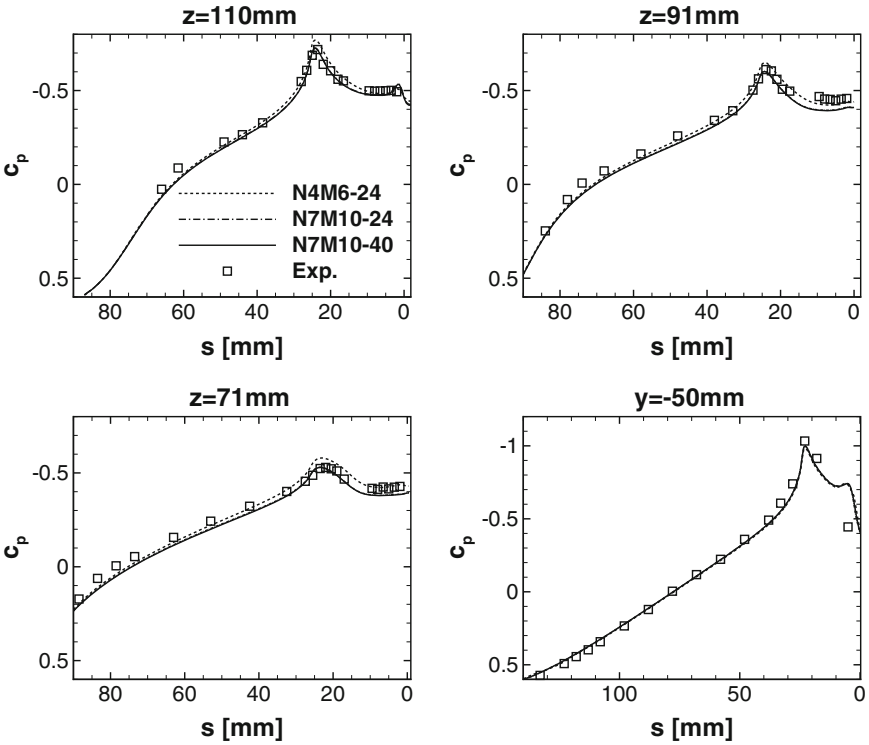


Fig. 25 Computational and measured pressure coefficient distributions along surface lines $z = 110$, $z = 71$ and $y = -100\,\text{mm}$. s denotes the wall-tangential distance to the trailing edge

The wall-normal and wall-tangential grid spacings Δy and Δx are given with respect to the inner-element resolution, which takes the degrees of freedom within each cell into account: $\Delta y = \Delta y_{Element}/(N+1)$. They represent maximum values in the refined region on the side surface. All simulations were conducted on the CRAY XC40 Hornet cluster at HLRS. The wall time per convective time $T^* = L/u_\infty$ on 3288 cores for simulations $N7M10$ and $N4M6$ amounted to about 4.6 and 0.6 h, respectively.

Figure 25 compares the time-averaged surface pressure coefficient for the two resolutions. The pressure coefficient is extracted along lines with $z = \text{const.}$ on the side surface and $y = \text{const.}$ on the upper surface. Experimental data from Werner et al. (2017a) is given for comparison. Additionally, for the case $N7M10$, two averaging periods (24 and 40 T^*) are compared. The results for the different averaging windows are not discernible, suggesting that the chosen time frames are sufficient and a statistically steady mean flow is reached. With regards to the p-refinement, it should be noted that the resolution for case $N7M10$ corresponds nearly to an isotropic doubling of the resolution, even without taking the more accurate approximation into account. Thus, the slight difference between the two resolution indicates that a

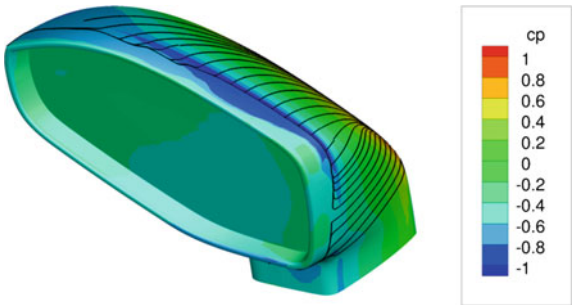


Fig. 26 Isocontours of the time-averaged pressure coefficient and surface streamlines based on wall friction

regime of weak grid dependence is reached. Based on these findings and taking into account the close agreement with the experimental static pressure measurements, the following analysis focuses on the highest resolution case.

Time-Averaged Flow Field The time-averaged flow field is characterized by the pressure coefficient on the mirror surface. Figure 26 shows the corresponding c_p distribution as well as the surface streamlines based on the skin friction on the leeward side. About 25 mm upstream of the trailing edge, shortly downstream of the design edge (marked "DE" in Fig. 23), the coalescing skin friction lines indicate a boundary layer separation, supported also by the increase in pressure along the trailing edge.

From the experiments, one possible source of tonal noise has been located at the side surface. To quantify the boundary layer in that area, a Particle Image Velocimetry (PIV) measurement campaign was conducted at the $z = 110$ mm position (see Fig. 23). The PIV data is plotted alongside the results from the numerical simulation. Figure 27 (left and right) shows contour plots of the time-averaged velocity magnitude and root mean square (RMS) velocity fluctuations. Overall, LES and PIV

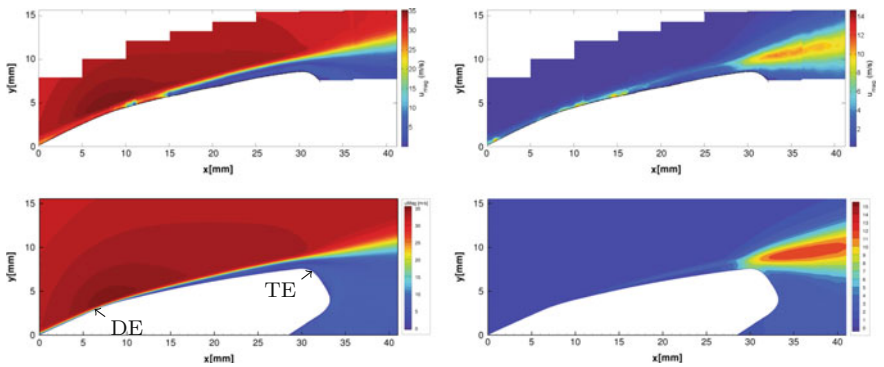


Fig. 27 Comparison of the simulation results (*bottom*) with PIV data (*top*) in the $z = 110$ mm plane. *Left* time-averaged velocity magnitude $\langle \vec{v} \cdot \vec{v} \rangle^{1/2}$, *Right* RMS velocity fluctuations $\langle \vec{v}' \cdot \vec{v}' \rangle^{1/2}$. The origin of the coordinate system in this plot is arbitrarily shifted to match the PIV data

data are in good agreement. The point of separation is predicted by the LES shortly downstream of the design edge in both cases, while the experimental data shows some artifacts in that region which can be attributed to seeding material deposition between the measured slice and the camera. The spreading rate of the shear layer and its separation angle are in very close agreement. From the RMS fluctuations, it can be determined that the flow remains laminar and steady through the separation up to the trailing edge, where the region of large amplitude fluctuations begins.

Acoustic Field and Source Identification Based on the comparisons of the hydrodynamic flow field with experimental data in the previous section, the focus is now shifted towards the acoustic emissions and source locations, with a focus on the occurrence and description of tonal noise. During the simulation, the local pressure signal is recorded at a rate of 44.1 kHz over $45T^*$ at 4000 position along a circular array of radius $r = 500$ mm. From this data, the PSD is computed using blocks with 2048 samples and 50% overlap. To reduce spectral leakage for non-periodic signals, a Hanning window is used.

A visual impression of the spatial distribution of the acoustic field is given in Fig. 28. Contours of the sound pressure level (SPL) of selected frequencies and the

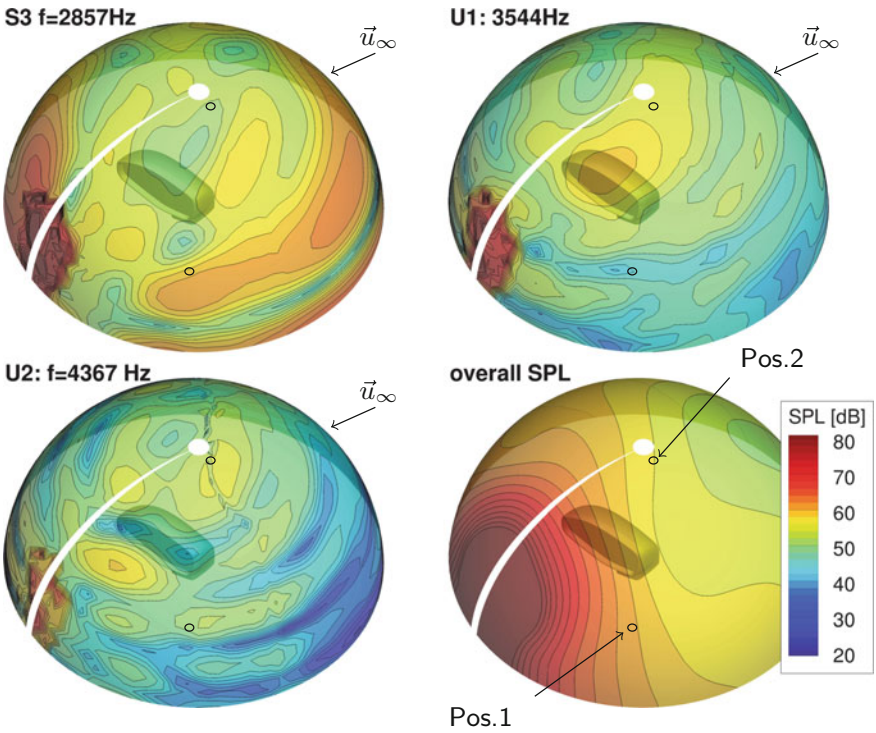


Fig. 28 SPL for selected frequencies and overall SPL (*bottom right*) on a spherical evaluation surface of $r = 500$ mm placed around the mirror

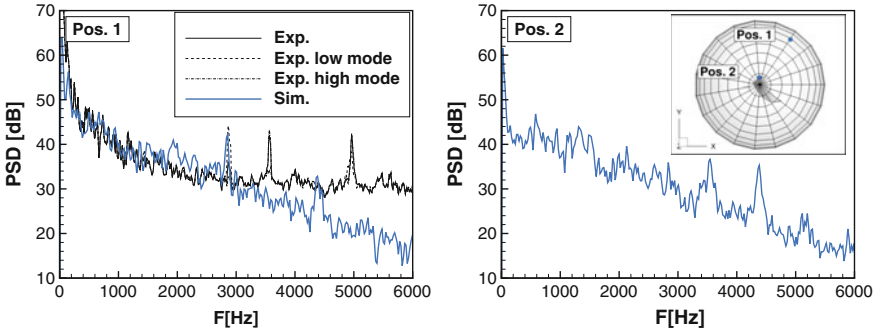


Fig. 29 PSD of pressure at two representative positions outside of the unsteady hydrodynamic field. PSD reference value: $4 \cdot 10^{-10}\,\mathrm{Pa}^2/\mathrm{Hz}$. The inset in the right panel shows the probe positions relative to the mirror geometry

overall SPL are plotted on a half-sphere above the mirror. As expected, the overall SPL increases significantly downstream of the mirror, due to the primary location of the noise sources at the leeward side of the mirror and the upstream shielding effects. Also, the exiting turbulent wake represents a strong source of noise. The acoustic footprint of the wake manifests itself as a "loud" spot across all selected frequencies. For each chosen frequency, a complex spatial wave pattern can be observed. For $f = 2857\,\mathrm{Hz}$, a strong lateral radiation can be observed, which also extends upstream. For $f = 3544$ and $4367\,\mathrm{Hz}$, more focused noise spots above and downstream of the mirror can be observed. Based on this qualitative analysis, we can expect that the frequency spectra vary significantly with the probe position. Therefore, the power spectral densities (PSD) of pressure in Fig. 29 are plotted at two representative probe positions, aiming at capturing the acoustic emission from the side surface at Pos. 1 and those from the upper surface at Pos. 2. The probes are located on a sphere with $r = 500\,\mathrm{mm}$ around the mirror, their positions are depicted in the right panel. The vertical positions are $z = 270$ and $z = 500\,\mathrm{mm}$ above the bottom wall for Pos. 1 and Pos. 2, respectively. The left panel includes inflow microphone measurements.

Before discussing the results, some remarks on the experimental setup and the comparability are necessary, which might help explain the results below. While in the simulation, a perfect free stream around the mirror is chosen, the experiments were conducted in a closed, rectangular test section. The scattering on the enclosure walls thus can be expected to influence the acoustic measurements. In addition to these effects, the transition locations and turbulence intensities in the parts of the flow around the mirror emitting the relevant broadband components cannot be guaranteed to match experimental ones, since the inflow turbulence level was not considered. Finally, the background noise of the wind tunnel is not captured in the simulation. Therefore, the following quantitative comparison focuses on the tonal noise frequencies, while the broadband noise spectra or amplitudes cannot be expected to match. Thus, the goal is the detection and comparison of the tonal components.

At Pos. 1, the simulated acoustic spectrum is composed of an evenly decaying broadband part and two recognizable, focused peaks at around 2860 Hz (S3) and 4380 Hz (U2). The first peak corresponds very well with the one found at 2900 Hz radiating from the mirror side during the experiments. However, an additional peak at 3500 Hz was also observed experimentally. A closer analysis revealed that the two modes alternated intermittently in an irregular fashion. Essentially, only one of them was noticeable at a given instant. The experimental spectrum is thus the result of averaging about the associated periods. Therefore, two additional experimental spectra corresponding to the low and high modes gained with conditional averaging are included in the plot. The simulation apparently predicts a situation where the lower of the two modes is favored. A switching of the modes was not observed numerically. While the precise reason for the alternation in the experiment is unknown, the switching between the two regimes triggered by loudspeaker forcing was shown in Werner et al. (2017b), indicating a high sensitivity to environmental disturbances of the flow in the experimental setup. Since the simulation setup is controlled and fixed, it is conceivable that the switching does not occur without voluntary triggering. In addition, due to the finite computational resources, the averaging time was significantly shorter than the observed alternation periods.

At Pos. 2, the computational data exhibits two tones at 3550 Hz (U1) and 4380 Hz (U2), which originate at the upper surface. The latter tone is also observed at Pos. 1, while the first is not, which can be explained by referring to Fig. 28: Pos. 1 lies within a shadowed region regarding the acoustic propagation of U1. The experimental spectrum at Pos. 1 only exhibits a single peak radiated from the upper surface, which has a significantly higher frequency of about 5000 Hz (Fig. 29 (left)). An indication for a tone of a similar frequency in the simulation is found in the weak trace of a peak at about 5000 Hz in the computational data at Pos. 2, which is generated at the upper surface.

In order to locate the dominant noise sources, an experimental beam-forming with a linear microphone array was conducted. Results indicate that the main source regions are located around the airfoil trailing edge, analogously to airfoil self-noise. Therefore, in Fig. 30, the spectra of wall pressure fluctuations along the trailing edge are plotted. The local coordinate r traverses the trailing edge from the lower side surface to the outer top surface. The side and upper portion are marked in Fig. 30. The PSD spectra are calculated using using blocks of 1024 samples averaged over $28T^*$. On both the side and the upper surface, two areas are marked that contain clear narrowband frequency peaks. Specifically, the side surface features the expected peak at approx. 2860 Hz corresponding to S3, while on the upper surface multiple additional narrowband features are visible. Among these features we recognize U1, U2 and the weak trace at 5000 Hz. Thus, each tone observed in the acoustic spectra has a counterpart in the wall pressure spectrum. The various tonal noise components therefore originate from the respective hydrodynamic fluctuations at the trailing edge.

Unsteady Flow Field In the previous section, the wall pressure fluctuations were connected to the generation of tonal noise. In order to characterize the underlying unsteady hydrodynamic field, its instantaneous vortical structures are visualized in

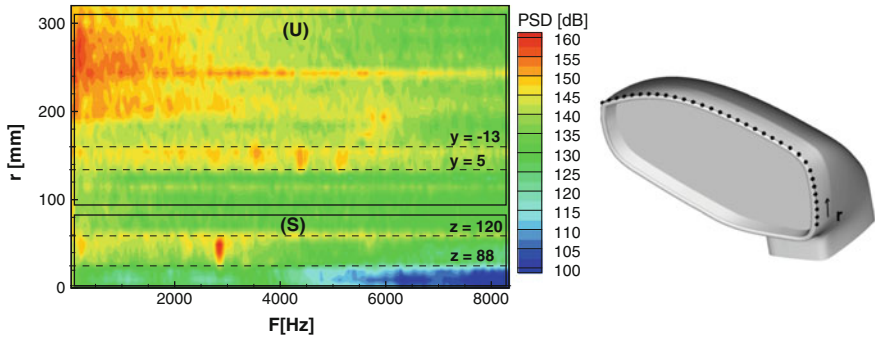


Fig. 30 Power spectral density of the wall pressure along the circumferential coordinate r at the trailing edge. PSD reference value: $4 \cdot 10^{-10}/\mathrm{Pa}^2\,\mathrm{Hz}$

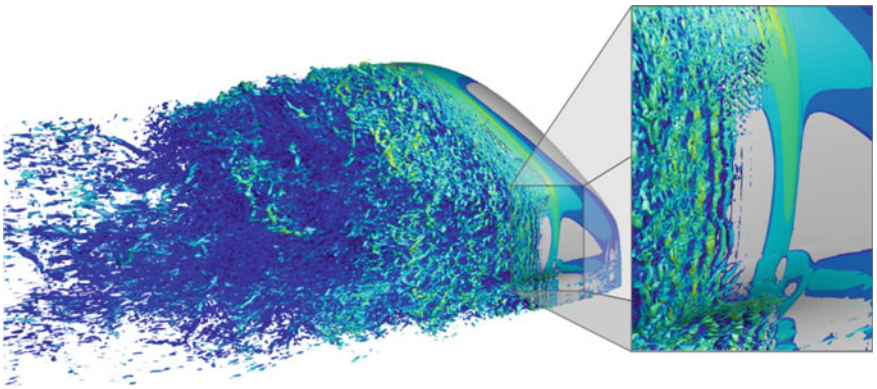


Fig. 31 Vortical structures at a flow field snapshot visualized by isosurfaces of $Q = 100(u_\infty/L)^2$ colored with velocity magnitude

Fig. 31 by the means of isosurfaces of the Q-criterion (Haller 2005), colored by velocity magnitude.

The initial laminar flow appears as a smooth surface on the front side of the mirror. The first vortical structures appear between the design edge and the trailing edge. On the side surface, a regular pattern of spanwise oriented rollers emerge. To evaluate the frequency associated with these coherent structures, which are a clear candidate for the development of tonal noise, a discrete temporal Fourier transform is performed. Details on the specific analysis can be found in Frank and Munz (2016), Frank (2016) and strongly support the notion that the tone associated with the side surface S3 originates from the passing of these structures over the trailing edge. A visual impression of the associated spatial structures is given in Fig. 32, where tonal mode S3 is compared to two representative modes of the surrounding broadband range. Shown are the isosurfaces of streamwise velocity fluctuations in the top row, and the isocontours of pressure fluctuations in the $z = 110\,\mathrm{mm}$ cut in the bottom

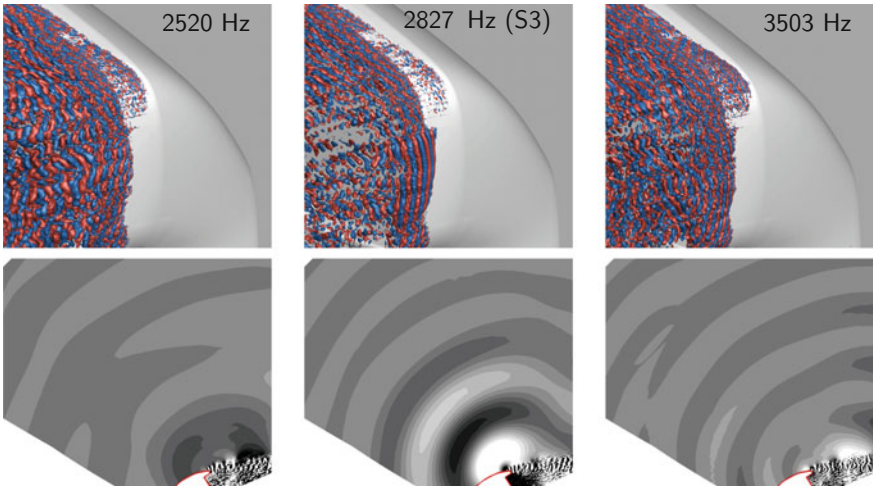


Fig. 32 *Top* isosurfaces of (the real part of) positive and negative velocity in the streamwise direction. The levels are chosen to ensure comparability. *Bottom* pressure contours at $z = 110\,\text{mm}$. The *left*, *middle* and *right columns* correspond to $f = 2520$, 2827 and 3503 Hz

row. Note that plotting the real part results in an arbitrary but spatially consistent phase. The isosurfaces of S3 exhibit clear ordered coherent structures on the side surface, and the associated acoustic field indicates a high amplitude tonal source in the direct vicinity of the trailing edge. For the other two frequencies, such clear levels of coherence cannot be identified, and no clear statement regarding the source position can be made.

In summary, in this section we have demonstrated how the DGSEM framework can be used successfully in the direct noise computation in challenging domains. We have shown a near perfect comparison of the hydrodynamic field to the experimental data, and a very close agreement between the simulated and measured emitted noise frequencies. A global stability analysis similar to the one presented in Sect. 4.2 was conducted which confirmed the existence of the feedback loop and showed very good agreement of the loop frequency with the phase condition. Further details can be found in Frank and Munz (2016). To the authors' knowledge, this constitutes the first numerical simulation of the tonal feedback mechanism at a three-dimensional, complex geometry.

5 Conclusion

In this chapter, we have given an overview of the state of the art of direct acoustic simulation with Discontinuous Galerkin methods. Many variants of DG methods exist, which mainly differ in implementation details, meshing flexibility and

computational efficiency. However, they all share the basic advantages of the method for DNS and DNC: They allow arbitrary order in space, which supports excellent wave propagation properties and thus reduces the number of degrees of freedom required to simultaneously resolve small scale fluctuations alongside large scale structures. Due to the inter-element numerical fluxes, they are also naturally suited for hyperbolic problems and thus are an attractive base scheme for multi-scale problems such as the acoustic noise generation and emission arising from the compressible Navier–Stokes equations. Since the boundary conditions can be enforced weakly through characteristic-splitting based flux functions, far-field acoustic boundary conditions can be applied in a straight-forward manner. For the outflow boundary, where large amplitude nonlinear structures exit, an absorbing layer approach is feasible. We have presented such a sponge zone approach based on an adaptive, temporally filtered base flow, which has the advantage of preserving the time-averaged hydrodynamic flow field.

The framework FLEXI is based on a specific, highly efficient variant of the DG family. It has shown excellent scaling on high performance computing clusters for large scale simulations of turbulence. With the help of recent additions in terms of boundary conditions and analysis postprocessing tools, FLEXI has been extended towards challenging direct noise computations in complex domains. We have demonstrated the suitability of the framework for DNC, in particular for the exploratory numerical investigations into complex interactions of noise and flow such as the feedback mechanism. The numerical simulation of this mechanism, which has been identified as a main source of tonal noise around bluff bodies, demands a highly accurate numerical scheme for laminar, transitional and turbulent regions of the flow as well as a faithful resolution of the geometry. With the help of the DG methods, the numerical representation of this feedback mechanism at a complex automotive side mirror was possible for the first time.

References

Akervik, E., Brandt, L., Henningson, D. S., Hoepffner, J., Marxen, O., & Schlatter, P. (2006). Steady solutions of the Navier-Stokes equations by selective frequency damping. *Physics of Fluids*, *18*(6), 068102.

Arbey, H., & Bataille, J. (1983). Noise generated by airfoil profiles placed in a uniform laminar flow. *Journal of Fluid Mechanics*, *134*, 33–47. ISSN 1469-7645, 0022-1120.

Bassi, F., & Rebay, S. (1997). A high-order accurate discontinuous finite element method for the numerical solution of the compressible Navier-Stokes equations. *Journal of Computational Physics*, *131*(2), 267–279. ISSN 0021-9991.

Bazilevs, Y., & Hughes, T. J. R. (2007). Weak imposition of Dirichlet boundary conditions in fluid mechanics. *Computers and Fluids*, *36*(1), 12–26. ISSN 0045-7930.

Beck, A. (2015). High order discontinuous Galerkin methods for the simulation of multiscale problems. Ph.D. thesis, University of Stuttgart.

Beck, A. D., Bolemann, T., Flad, D., Frank, H., Gassner, G., Hindenlang, F., et al. (2014). High-order discontinuous Galerkin spectral element methods for transitional and turbulent flow simulations. *International Journal for Numerical Methods in Fluids*, *76*(8), 522–548.

Beck, A. D., Flad, D. G., Tonhäuser, C., Gassner, G., & Munz, C.-D. (2016). On the influence of polynomial de-aliasing on subgrid scale models. *Flow, Turbulence and Combustion*, 1–37.

Bogey, C., & Bailly, C. (2004). A family of low dispersive and low dissipative explicit schemes for flow and noise computations. *Journal of Computational Physics*, *194*(1), 194–214. ISSN 0021-9991.

Burstedde, C., Wilcox, L., & Ghattas, O. (2011). p4est: Scalable algorithms for parallel adaptive mesh refinement on forests of octrees. *SIAM Journal on Scientific Computing*, *33*(3), 1103–1133. ISSN 1064-8275.

Choi, H., & Moin, P. (2012). Grid-point requirements for large eddy simulation: Chapman's estimates revisited. *Physics of Fluids*, *24*, 011702–011702. ISSN 0899-8213.

Chong, T. P., & Joseph, P. (2012). Ladder- structure in tonal noise generated by laminar flow around an airfoil. *The Journal of the Acoustical Society of America*, *131*(6), EL461–EL467. ISSN 0001-4966.

Cockburn, B., & Shu, C.-W. (1989). TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws II: General framework. *Mathematics of Computation*, *52*(186), 411–435. ISSN 0025-5718.

Cockburn, B., & Shu, C.-W. (1991). The Runge-Kutta local projection p1-discontinuous-Galerkin finite element method for scalar conservation laws. *ESAIM: Mathematical Modelling and Numerical Analysis*, *25*(3), 337–361. ISSN 0764-583X, 1290-3841.

Cockburn, B., Lin, S.-Y., & Shu, C.-W. (1989). TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws III: One-dimensional systems. *Journal of Computational Physics*, *84*(1), 90–113. ISSN 0021-9991.

Cockburn, B., Hou, S., & Shu, C. W. (1990). The Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. IV: The multidimensional case. *Mathematics of Computation*, *54*(190), 545–581. ISSN 0025-5718.

Collis, S. S. (2002). Discontinuous Galerkin methods for turbulence simulation. In *Proceedings of the 2002 Center for Turbulence Research Summer Program* (pp. 155–167).

Colonius, T. (2004). Modeling artificial boundary conditions for compressible flow. *Annual Review of Fluid Mechanics*, *36*, 315–345.

Colonius, T., & Lele, S. K. (2004). Computational aeroacoustics: Progress on nonlinear problems of sound generation. *Progress in Aerospace Sciences*, *40*(6), 345–416. ISSN 0376-0421.

Desquesnes, G., Terracol, M., & Sagaut, P. (2007). Numerical investigation of the tone noise mechanism over laminar airfoils. *Journal of Fluid Mechanics*, *591*, 155–182. ISSN 1469-7645, 0022-1120.

Fechter, S., Hindenlang, F., Frank, H., Munz, C.-D., & Gassner, G. (2012) Discontinuous Galerkin schemes for the direct numerical simulation of fluid flow and acoustics. In *18th AIAA/CEAS Aeroacoustics Conference (33rd AIAA Aeroacoustics Conference)*. American Institute of Aeronautics and Astronautics. doi:10.2514/6.2012-2187.

Flad, D., Beck, A. D., Gassner, G., & Munz, C.-D. (2014). A discontinuous Galerkin spectral element method for the direct numerical simulation of aeroacoustics. In *20th AIAA/CEAS Aeroacoustics Conference*. American Institute of Aeronautics and Astronautics. doi:10.2514/6.2014-2740.

Frank, H. M. (2016). High order large eddy simulation for the analysis of tonal noise generation via aeroacoustic feedback effects at a side mirror. Ph.D. thesis, University of Stuttgart.

Frank, H. M., & Munz, C.-D. (2016). Direct aeroacoustic simulation of acoustic feedback phenomena on a side-view mirror. *Journal of Sound and Vibration*, *371*, 132–149. ISSN 0022-460X.

Gassner, G., & Kopriva, D. A. (2011). A comparison of the dispersion and dissipation errors of Gauss and Gauss-Lobatto discontinuous Galerkin spectral element methods. *SIAM Journal of Scientific Computing*, *33*(5), 2560–2579. ISSN 1064-8275.

Haller, G. (2005). An objective definition of a vortex. *Journal of Fluid Mechanics*, *525*, 1–26. ISSN 1469-7645, 0022-1120.

Hindenlang, F. (2014). Mesh curving techniques for high order parallel simulations on unstructured meshes. Ph.D. thesis, University of Stuttgart.

Hindenlang, F., Gassner, G., Altmann, C., Beck, A., Staudenmaier, M., & Munz, C.-D. (2012). Explicit discontinuous Galerkin methods for unsteady problems. *Computers and Fluids*, *61*, 86–93. ISSN 0045-7930.

Hindenlang, F., Bolemann, T., & Munz, C.-D. (2015). Mesh curving techniques for high order discontinuous Galerkin simulations. In N. Kroll, C. Hirsch, F. Bassi, C. Johnston, & K. Hillewaert (Eds.), *IDIHOM: Industrialization of high-order methods - a top-down approach* (Vol. 128, pp. 133–152). Notes on numerical fluid mechanics and multidisciplinary design. New York: Springer International Publishing. doi:10.1007/978-3-319-12886-3_8. ISBN 978-3-319-12885-6 978-3-319-12886-3.

Howe, M. S. (2003). *Theory of vortex sound*. Cambridge: Cambridge University Press.

Hussaini, M. Y., Kopriva, D. A., Salas, M. D., & Zang, T. A. (1985). Spectral methods for the Euler equations. i - Fourier methods and shock capturing. *AIAA Journal*, *23*(1), 64–70. ISSN 0001-1452.

Jones, L., & Sandberg, R. (2010). Numerical investigation of tonal airfoil self-noise generated by an acoustic feedback-loop. In *16th AIAA/CEAS Aeroacoustics Conference*. American Institute of Aeronautics and Astronautics. doi:10.2514/6.2010-3701.

Jones, L. E., & Sandberg, R. D. (2011). Numerical analysis of tonal airfoil self-noise and acoustic feedback-loops. *Journal of Sound and Vibration*, *330*(25), 6137–6152. ISSN 0022-460X.

Khalighi, Y., Mani, A., Ham, F., & Moin, P. (2010). Prediction of sound generated by complex flows at low mach numbers. *AIAA Journal*, *48*(2), 306–316. ISSN 0001-1452.

Kingan, M. J., & Pearse, J. R. (2009). Laminar boundary layer instability noise produced by an aerofoil. *Journal of Sound and Vibration*, *322*(4), 808–828. ISSN 0022-460X.

Kirby, R. M., & Karniadakis, G. E. (2003). De-aliasing on non-uniform grids: Algorithms and applications. *Journal of Computational Physics*, *191*(1), 249–264. ISSN 0021-9991.

Kolmogorov, A. N. (1999). The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers. *Royal Society of London Proceedings Series A*, *434*, 9–13. doi:10.1098/rspa.1991.0075.

Kopriva, D. A. (2006). Metric identities and the discontinuous spectral element method on curvilinear meshes. *Journal of Scientific Computing*, *26*(3), 301. ISSN 0885-7474, 1573-7691.

Kopriva, D. A. (2009). *Implementing spectral methods for partial differential equations: Algorithms for scientists and engineers* (1st ed.). New York: Springer Publishing Company Incorporated. ISBN 9048122600, 9789048122608.

Kopriva, D. A., & Gassner, G. (2010). On the quadrature and weak form choices in collocation type discontinuous Galerkin spectral element methods. *Journal of Scientific Computing*, *44*(2), 136–155. ISSN 0885-7474, 1573-7691.

Lele, S. K. (1997). Computational aeroacoustics: A review. *AIAA Paper*, *18*, 1997.

Lesaint, P., & Raviart, P.-A. (1974). On a finite element method for solving the neutron transport equation. In C. A. deBoor (Ed.), *Mathematical aspects of finite elements in partial differential equations* (pp. 89–145). New York: Academic Press.

Lighthill, M. J. (1952). On sound generated aerodynamically. i. General theory. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, *211*(1107), 564–587. ISSN 1364-5021, 1471-2946.

Lounsberry, T. H., Gleason, M. E., & Puskarz, M. M. (2007). Laminar flow whistle on a vehicle side mirror. In SAE Technical Paper. SAE International, 04.

Nash, E. C., Lowson, M. V., & McAlpine, A. (1999). Boundary-layer instability noise on aerofoils. *Journal of Fluid Mechanics*, *382*, 27–61. ISSN 1469-7645, 0022-1120.

Paterson, R. W., Vogt, P. G., Fink, M. R., & Munch, C. L. (1973). Vortex noise of isolated airfoils. *Journal of Aircraft*, *10*(5), 296–302. ISSN 0021-8669.

Plogmann, B., Herrig, A., & Wuerz, W. (2013). Experimental investigations of a trailing edge noise feedback mechanism on a NACA 0012 airfoil. *Experiments in Fluids*, *54*(5), 1480. ISSN 0723-4864, 1432-1114.

Pruett, C. D., Gatski, T. B., Grosch, C. E., & Thacker, W. D. (2003). The temporally filtered Navier-Stokes equations: Properties of the residual stress. *Physics of Fluids*, *15*(8), 2127–2140. ISSN 1070-6631.

Reed, W. H., & Hill, T. R. (1973). Triangular mesh methods for the neutron transport equation. Technical report LA-UR-73-479, Los Alamos Scientific Laboratory.

Shebalin, J. (1993). Pseudospectral simulation of compressible turbulence using logarithmic variables. In *11th Computational Fluid Dynamics Conference*. American Institute of Aeronautics and Astronautics.

Tam, C., & Webb, J. C. (1993). Dispersion-relation-preserving finite difference schemes for computational acoustics. *Journal of Computational Physics*, *107*(2), 262–281. ISSN 0021-9991.

Tam, C. K. W. (1974). Discrete tones of isolated airfoils. *Journal of the Acoustical Society of America*, *55*, 1173–1177.

Tam, C. K. W., & Hardin, J. C. (1997). Second computational aeroacoustics (CAA) workshop on benchmark problems. NASA Conference Publications.

Toro, E. F. (1999). *Riemann solvers and numerical methods for fluid dynamics*. New York: Springer.

Werner, M. (2017). Experimental study on tonal self-noise generated by aeroacoustic feedback on a side mirror. Ph.D. thesis, University of Stuttgart.

Werner, M. J., Würz, W., & Krämer, E. (2017a) Experimental investigation of an aeroacoustic feedback mechanism on a two-dimensional side mirror model. *Journal of Sound and Vibration*, *387*, 79–95. ISSN 0022-460X.

Werner, M. J., Würz, W., & Krämer, E. (2017b). Experimental investigation of the tonal self-noise emission of a vehicle side mirror. *Accepted for publication in AIAA Journal*.

Yokokawa, M., Itakura, K., Uno, A., Ishihara, T., & Kaneda, Y. (2002). 16.4-tflops direct numerical simulation of turbulence by a Fourier spectral method on the earth simulator. In *Supercomputing, ACM/IEEE 2002 Conference* (pp. 50–50).

Direct and Iterative Solvers

Ulrich Langer and Martin Neumüller

Abstract This chapter on solvers gives a compact introduction to direct and iterative solvers for systems of algebraic equations typically arising from the finite element discretization of partial differential equations or systems of partial differential equations. Beside classical iterative solvers, we also consider advanced preconditioning and solving techniques like additive and multiplicative Schwarz methods, generalizing Jacobi's and Gauss–Seidel's ideas to more general subspace correction methods. In particular, we consider multilevel diagonal scaling and multigrid methods.

1 Introduction

We will start our chapter on SOLVERS with a section about the systems of linear algebraic equations typically arising in Computational Acoustics and their properties that are important for the behavior of the classical direct and iterative solvers. The second section is devoted to direct solvers, where we discuss the standard algorithms like the classical Gauss algorithm, the LU decomposition and direct solvers utilizing the sparsity of the system matrix including direct solvers that are nowadays called sparse direct solvers. In the next section about iterative solvers, we introduce the classical iterative methods like the Jacobi method, the Gauss–Seidel method and the Richardson method. We will also focus on more advanced iterative methods like gradient and Conjugate Gradient (CG) methods for linear systems with Symmetric and Positive Definite (SPD) system matrices. The CG method is the most prominent method from the general class of Krylov subspace methods. The Generalized

U. Langer (✉)
Johann Radon Institute for Computational and Applied Mathematics,
Austrian Academy of Sciences, Linz, Austria
e-mail: ulanger@numa.uni-linz.at

U. Langer · M. Neumüller
Institute of Computational Mathematics, Johannes Kepler University, Linz, Austria

M. Kaltenbacher (ed.), *Computational Acoustics*, CISM International Centre
for Mechanical Sciences 579, DOI 10.1007/978-3-319-59038-7_5

Minimal Residual (GMRES) method is another famous representative of Krylov subspace methods that can be used for solving algebraic systems with general regular system matrices. Since the convergence behaviour of all these iterative methods heavily depends on the conditioning of the underlying linear system, we will provide a section on preconditioning techniques. The final technique which will be explained in this section are subspace correction methods. We will close our contribution with an introduction to multigrid methods. We motivate this section by again looking at classical iterative schemes and their properties. By combining the iterative schemes with subspace correction methods (coming from the previous section), we derive multigrid methods and explain the main aspects of these methods and their implementation. We conclude this section by introducing time-multigrid methods which are suited for parallelization with respect to time.

2 Linear Systems of Algebraic Equations

The efficient solution of linear systems is a fundamental task in all Computational Sciences and, in particular, in Computational Acoustics (CA). These systems are typically arising from the discretization and possible linearization of Partial Differential Equations (PDEs) or systems of coupled PDEs. Thus, these systems have special structures and properties which can be used when one constructs efficient solution techniques. The problem of solving such systems can typically be posed in the following form: Given some regular $n_h \times n_h$ system matrix $\mathbf{A} = \left[A_{ij}\right]_{i,j=1,\dots,n_h} \in \mathbb{R}^{n_h \times n_h}$ and some right-hand side (rhs) vector $\underline{b} = [b_i]_{i=1,\dots,n_h} \in \mathbb{R}^{n_h}$, find the solution vector $\underline{u} = \left[u_j\right]_{j=1,\dots,n_h} \in \mathbb{R}^{n_h}$ of the system

$$\mathbf{A}\underline{u} = \underline{b}, \tag{1}$$

where $n = n_h = n_{eq} = \mathcal{O}(h^{-d})$ is the number of unknowns (degree of freedoms = dofs) that is equal to the number of equations in (1), h denotes the discretization parameter, and d is the space dimension of the computational domain $\Omega \subset \mathbb{R}^d$ in which the PDEs are posed.

Possible system matrices that are typically arising in Computational Acoustics are the following:
$\mathbf{A} = \mathbf{D} = \text{diag}(D_{ii})$ - diagonal matrix (mass lumping),
$\mathbf{A} = \mathbf{M}$ - mass matrix,
$\mathbf{A} = \mathbf{K}$ - stiffness matrix,
$\mathbf{A} = \mathbf{M} + \gamma_H\, \Delta t\, \mathbf{C} + \beta_H (\Delta t)^2 \mathbf{K}$ - Newmark matrix,
$\mathbf{A} = \mathbf{K} - \omega^2 \mathbf{M}$ - time-harmonic case,
$\mathbf{A} = \mathbf{B}$ - fully populated BEM matrices,
where $\mathbf{C}$ here denotes some damping matrix.

Let us now consider the mixed Boundary Value Problem (BVP) for the potential equation as typical model problem: Given functions f, q_n, u_e and ν, find the unknown potential function u such that

$$-\nabla\cdot(\nu\nabla u) = f \text{ in } \Omega, \quad u = u_e := 0 \text{ on } \Gamma_e, \quad \frac{\partial u}{\partial \mathbf{n}} := \nabla u\cdot\mathbf{n} = q_n \text{ on } \Gamma_n, \quad (2)$$

where Γ_e and Γ_n are non-intersecting parts of the boundary $\Gamma = \partial\Omega$ of the computational domain Ω on which Dirichlet (essential) and Neumann (natural) boundary conditions are prescribed, respectively. The computational domain $\Omega \subset \mathbb{R}^d$ is assumed to be bounded and Lipschitz. The dimension d is usually equal to 1, 2, or 3. In (2), ∇ and $\mathbf{n}$ denote the gradient and the outer normal to Ω, respectively. In the case $\nu = 1$, the PDE (2) becomes the famous Poisson equation $-\Delta u = f$. The weak formulation of (2) reads as follows: Find $u \in V_{u_e} := \{v \in H^1(\Omega) : v = u_e \text{ on } \Gamma_e\}$ such that

$$\int_\Omega \nu(\mathbf{x})\nabla u(\mathbf{x})\cdot\nabla v(\mathbf{x})\,d\mathbf{x} = \int_\Omega f(\mathbf{x})v(\mathbf{x})d\mathbf{x} + \int_{\Gamma_n} q_n(\mathbf{x})v(\mathbf{x})\,ds \quad (3)$$

for all test functions $v \in V_0 := \{v \in H^1(\Omega) : v = 0 \text{ on } \Gamma_e\}$, where

$$H^1(\Omega) = \{v \in L_2(\Omega) : \text{ there exists the weak gradient } \nabla v \in L_2(\Omega)\}$$

denotes the Sobolev space $W_2^1(\Omega)$ that is equipped with the norm

$$\|v\|_1 := \sqrt{\|v\|_0^2 + |v|_1^2} = \sqrt{\int_\Omega |v|^2 d\mathbf{x} + \int_\Omega |\nabla v|^2 d\mathbf{x}}.$$

The weak or variational formulation (3) can be written in the more abstract variational form: Find $u \in V_{u_e}$ such that

$$a(u, v) = \ell(v) \quad \forall v \in V_0, \quad (4)$$

where the bilinear form $a(\cdot,\cdot)$ and the linear form $\ell(\cdot)$ are given by the left-hand and right-hand sides of (3), respectively. This is indeed a very general setting of the variational formulation for elliptic boundary value problems.

In order to investigate existence and uniqueness of a weak solution of (3), we can assume that the Dirichlet data $u_e = 0$, otherwise we assume that u_e is a trace of some function u_e (we use the same notation) from $H^1(\Omega) \cap C(\overline{\Omega})$ on Γ_e, and make the ansatz $u = u_e + w$ where the unknown function w now fulfills homogeneous Dirichlet boundary conditions on Γ_e. This procedure is called homogenization. Furthermore, we assume that the given source function $f \in L_2(\Omega)$, the given Neumann data $q_n \in L_2(\Gamma_n)$, and ν is a given uniformly positive and bounded coefficient function, i.e. there are positive constants ν_1 and ν_2 such that

$$0 < \nu_1 \leq \nu(\mathbf{x}) \leq \nu_2 \tag{5}$$

for almost all $\mathbf{x} \in \Omega$.

The Lax–Milgram lemma (see, e.g., Steinbach 2008; Jung and Langer 2013) delivers existence (∃) and uniqueness (!) provided that the following assumptions with respect to the Hilbert space V_0 are fulfilled:

1. the linear form $\ell(\cdot)$ is nothing but a continuous (bounded), linear functional on V_0, i.e., there exists a non-negative constant c_ℓ such that

$$|\ell(v)| \leq c_\ell \, \|v\|_1, \ \forall v \in V_0, \tag{6}$$

2. the bilinear form $a(\cdot, \cdot)$ is continuous (bounded) on V_0, i.e., there exists a positive constant μ_2 such that

$$|a(u, v)| \leq \mu_2 \, \|u\|_1 \|v\|_1, \ \forall u, v \in V_0, \tag{7}$$

3. the bilinear form $a(\cdot, \cdot)$ is elliptic (coercive) on V_0, i.e., there exists a positive constant μ_1 such that

$$a(v, v) \geq \mu_1 \, \|v\|_1^2, \ \forall v \in V_0. \tag{8}$$

Using Cauchy's inequality and the trace inequality $\|v\|_{L_2(\Gamma_n)} \leq c_T(\Gamma_n)\|v\|_1$ that holds for all $v \in H^1(\Omega)$ (see, e.g., Steinbach 2008; Jung and Langer 2013), we easily see that (6) holds for our model problem (3) with $c_\ell = \|f\|_0 + c_T \|q_n\|_{L_2(\Gamma_n)}$. Furthermore, assumption (5) and Cauchy's inequality give $\mu_2 = \nu_2$. In order to prove V_0-ellipticity of the bilinear form $a(\cdot, \cdot)$, we need Friedrichs' inequality

$$\|v\|_0 \leq c_F(\Gamma_e) \, |v|_1, \ \forall v \in V_0, \tag{9}$$

see, e.g., Steinbach (2008), Jung and Langer (2013). Now, we can estimate as follows:

$$a(v, v) \geq \nu_1 |v|_1^2 = \nu_1 \frac{1}{2}(|v|_1^2 + |v|_1^2) \geq \nu_1 \frac{1}{2} \min\{c_F^{-2}, 1\} \|v\|_1^2 \ \forall v \in V_0,$$

i.e., $\mu_1 = \nu_1 \frac{1}{2} \min\{c_F^{-2}, 1\}$. Therefore, our model problem (3) has a unique weak solution.

The Finite Element Scheme is nothing than the standard Galerkin scheme, and can be written in the abstract form: Find $u^h \in V_{u_e}^h$ such that

$$a(u^h, v^h) = \ell(v^h) \quad \forall v^h \in V_0^h, \tag{10}$$

where the Finite Element test space $V_0^h := \text{span}\{N_1, N_2, \ldots, N_{n_{eq}}\} \subset V_0$ is spanned by all basis functions that vanish on Γ_e, whereas the manifold $V_{u_e}^h = u_e^h + V_0^h$, in which we look for the solution, is given by all functions of the form

$$\sum_{j=n_{eq}+1}^{n_n} u_e(\mathbf{x}_j) N_j(\mathbf{x}) + \sum_{j=1}^{n_{eq}} v_j N_j(\mathbf{x})$$

interpolating the Dirichlet data u_e at the nodes x_j located on Γ_n. Here, the superscript h is nothing but the usual finite element discretization parameter. Therefore, for our model problem (3), the Finite Element scheme (10) can be rewritten in the form: Find $u^h(\mathbf{x}) = \sum_{j=1}^{n_{eq}} u_j N_j(\mathbf{x}) + \sum_{j=n_{eq}+1}^{n_n} u_e(\mathbf{x}_j) N_j(\mathbf{x}) \in V_{u_e}^h$ such that

$$\int_\Omega \nu \nabla u^h \cdot \nabla v^h \, d\mathbf{x} = \int_\Omega f v^h d\mathbf{x} + \int_{\Gamma_n} q_n v^h \, ds \tag{11}$$

for all $v^h \in V_0^h := \text{span}\{N_1, N_2, \ldots, N_{n_{eq}}\}$. Now, already the ellipticity of the bilinear form $a(\cdot,\cdot)$ on $V_0^h \subset V_0$ implies that there exist a unique finite element solution $u^h \in V_{u_e}^h$ of (10) respectively (11). A priori and a posteriori estimates of the discretization error $u - u^h$ in different norms can be found in Sect. 4.5.4 in Jung and Langer (2013), or in Chap. 11 in Steinbach (2008).

Once the FE basis is chosen, the FE scheme (11) is obviously equivalent to the solution of a linear system of equations: Find $\underline{u} = \left[u_j\right]_{j=1,\ldots,n_h} \in \mathbb{R}^{n_h = n_{eq}}$ such that

$$\mathbf{K}\underline{u} = \underline{f}, \tag{12}$$

where $\mathbf{K} = \left[K_{ij}\right]_{i,j=1,\ldots,n_h}$ with $K_{ij} = \int_\Omega \nu \nabla N_j \cdot \nabla N_i \, d\mathbf{x}$, and $\underline{f} = [f_i]_{i=1,\ldots,n_h}$ with $f_i = \int_\Omega f N_i d\mathbf{x} + \int_{\Gamma_n} q_n N_i \, ds - \sum_{j=n_{eq}+1}^{n_n} K_{ij} u_e(\mathbf{x}_j)$. The assembling of the system matrix (stiffness matrix) $\mathbf{K}$ and the right-hand side (load vector) $\underline{f}$ is described in Sect. 4.5 of Jung and Langer (2013) in detail.

The stiffness matrix $\mathbf{K}$ obviously has the following structural properties:

- Large scale: $n_h = \mathcal{O}(h^{-d}) = 10^6$,…,$10^9$ dofs in practice!
- Sparse: $K_{ij} = 0 \, \forall i, j : \ \text{supp} N_i \cap \text{supp} N_j = \varnothing$, i.e., NNE = Number of Non-zero Elements $= \mathcal{O}(h^{-d}) = n_h$,
- Band respectively profile structure, i.e., $K_{ij} = 0$ if $|i - j| > b_w$ = bandwidth $= \mathcal{O}(h^{-(d-1)})$, where the band profile depends on the numbering of the nodes! There are heuristic algorithms of band or profile optimization like Cuthill–McKee algorithm, Reverse Cuthill–McKee algorithm, Minimal degree algorithm, see, e.g., Algorithms 5.16–5.18 in Jung and Langer (2013).

Furthermore, due to the so-called heredity relation

$$(\mathbf{K}\underline{u}, \underline{v}) := (\mathbf{K}\underline{u}, \underline{v})_{\mathbb{R}^n} = a(u^h, v^h) \ \forall \underline{u}, \underline{v} \leftrightarrow u^h, v^h \in V_0^h, \tag{13}$$

the stiffness matrix $\mathbf{K}$ inherits the properties of the bilinear form $a(\cdot,\cdot)$:

1. $a(u^h, v^h) = a(v^h, u^h) \ \forall u^h, v^h \in V_0^h \Rightarrow \mathbf{K} = \mathbf{K}^\top$.
2. $a(v^h, v^h) > 0 \ \forall v^h \in V_0^h \setminus \{0\} \Rightarrow \mathbf{K}$ is positive definite.

3. Thus, for our model problem (3), the stiffness matrix $\mathbf{K} = \mathbf{K}^\top > 0$ is Symmetric and Positive Definite (SPD), since $a(\cdot,\cdot)$ is symmetric and even V_0-elliptic.

Let us assume that $a(\cdot,\cdot)$ is symmetric, V_0-elliptic and V_0-bounded as it was proved for our model problem (3). Then we immediately get the following results:

1. $\mathbf{K}$ is SPD.
2. $\mathbf{K}$ has $n = n_h$ positive real eigenvalues λ_k with the corresponding eigenvectors $\underline{\varphi}_k$: $\mathbf{K}\underline{\varphi}_k = \lambda_k \underline{\varphi}_k$,

$$\begin{aligned} 0 < &\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n, \\ &\underline{\varphi}_1, \quad \underline{\varphi}_2, \quad \cdots \quad \underline{\varphi}_n, \end{aligned}$$

   where the eigenvectors can be chosen to be orthonormal, i.e.,

$$(\underline{\varphi}_i, \underline{\varphi}_j) := (\underline{\varphi}_i, \underline{\varphi}_j)_{\mathbb{R}^n} = \delta_{i,j} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{else} \end{cases} \tag{14}$$

   with respect to the Euclidian inner product in $\mathbb{R}^n$.
3. The spectral condition number is given by the relation

$$\kappa_2(\mathbf{K}) := \|\mathbf{K}\|\|\mathbf{K}^{-1}\| = \frac{\lambda_n}{\lambda_1} = \frac{\lambda_{\max}(\mathbf{K})}{\lambda_{\min}(\mathbf{K})}, \tag{15}$$

   where $\|\mathbf{K}\| = \max_{\underline{v}\in\mathbb{R}^n\setminus\underline{0}} \|\mathbf{K}\underline{v}\| \,/\, \|\underline{v}\|$ here denotes the spectral norm of $\mathbf{K}$, with $\|\underline{v}\|$ representing the Euclidian norm $\|\underline{v}\|_{\mathbb{R}^n} = (\underline{v}, \underline{v})^{1/2}_{\mathbb{R}^n}$.

In order to estimate the spectral condition number $\kappa_2(\mathbf{K})$, we use the representation of $\lambda_{\max}(\mathbf{K})$ and $\lambda_{\min}(\mathbf{K})$ by the maximum and minimum of the Rayleigh quotient $(\mathbf{K}\underline{v}, \underline{v})/(\underline{v}, \underline{v})$, respectively:

$$\lambda_{\max}(\mathbf{K}) = \max_{\underline{0}\neq\underline{v}\in\mathbb{R}^n} \frac{(\mathbf{K}\underline{v}, \underline{v})}{(\underline{v}, \underline{v})} \le c_2 h^{d-2} \tag{16}$$

and

$$\lambda_{\min}(\mathbf{K}) = \min_{\underline{0}\neq\underline{v}\in\mathbb{R}^n} \frac{(\mathbf{K}\underline{v}, \underline{v})}{(\underline{v}, \underline{v})} \ge c_1 h^{d}. \tag{17}$$

Estimate (16) can easily be derived from the heredity relation (13) and the computation of the maximal eigenvalues of the element stiffness matrices $\mathbf{K}^e$,

$$(\mathbf{K}\underline{v}, \underline{v}) = a(v^h, v^h) = \sum_{e=1}^{n_e} (\mathbf{K}^e \underline{v}^e, \underline{v}^e) \le \sum_{e=1}^{n_e} \lambda_{\max}(\mathbf{K}^e)(\underline{v}^e, \underline{v}^e),$$

whereas estimate (17) follows from

$$(\mathbf{K}\underline{v}, \underline{v}) = a(v^h, v^h) \geq \mu_1 \|v^h\|_1^2 \geq \mu_1 \|v^h\|_0^2 = \mu_1 (\mathbf{M}\underline{v}, \underline{v}) = \mu_1 \sum_{e=1}^{n_e} (\mathbf{M}^e \underline{v}^e, \underline{v}^e),$$

where $\mathbf{M}^e$ are the element mass matrices. The term $(\mathbf{M}^e \underline{v}^e, \underline{v}^e)$ in the sum can be estimated from below by $\lambda_{\min}(\mathbf{M}^e)(\underline{v}^e, \underline{v}^e)$ (Rayleigh quotient) that delivers the desired bound.

Estimates (16) and (17) immediately yield the spectral condition number estimate

$$\kappa_2(\mathbf{K}) = \frac{\lambda_{\max}(\mathbf{K})}{\lambda_{\min}(\mathbf{K})} \leq \frac{c_2}{c_1} h^{-2} \tag{18}$$

that is sharp with respect to (wrt) h, i.e., $\kappa_2(\mathbf{K}) = \mathcal{O}(h^{-2})$ for $h \to 0$, see also Example 2.1

Example 2.1 Let us consider the 1d example

$$-u''(x) = f(x), \ x \in (0,1), \quad u(0) = u(1) = 0 \tag{19}$$

yielding the FE stiffness matrix

$$\mathbf{K} = \frac{1}{h} \begin{pmatrix} 2 & -1 & 0 & \cdots & \cdots & 0 \\ -1 & 2 & -1 & \ddots & \ddots & \vdots \\ 0 & -1 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & -1 & 0 \\ \vdots & \ddots & 0 & -1 & 2 & -1 \\ 0 & \cdots & \cdots & 0 & -1 & 2 \end{pmatrix} \tag{20}$$

for hat functions $N_1, \ldots, N_{n_h=n-1}$ on a uniform grid $0 = x_0 < x_1 < \cdots < x_{n-1} < x_n = 1$ with $x_{i+1} - x_i = h = 1/n$. In this simple 1d case, the eigenvectors, the eigenvalues, and, therefore, the spectral condition numbers can be explicitly computed:

- Eigenvalues: $\lambda_k = \frac{4}{h} \sin^2(\frac{k\pi}{2n}), \ k = 1, 2, \ldots, n-1.$
- Eigenvectors: $\underline{\varphi}_k = [\sqrt{2n} \sin(k\pi i h)]_{i=1,\ldots,n-1}, k = 1, 2, \ldots, n-1.$
- Minimal eigenvalue:

$$\lambda_1 = \frac{4}{h} \sin^2\left(\frac{1\pi}{2n}\right) = \frac{4}{h} \sin^2\left(\frac{\pi h}{2}\right) = \mathcal{O}(h).$$

- Maximal eigenvalue:

$$\lambda_{n-1} = \frac{4}{h} \sin^2\left(\frac{(n-1)\pi}{2n}\right) = \frac{4}{h} \cos^2\left(\frac{\pi h}{2}\right) = \mathcal{O}(h^{-1}).$$

- Spectral condition number:

$$\kappa_2(\mathbf{K}) = \frac{\lambda_{\max}(\mathbf{K})}{\lambda_{\min}(\mathbf{K})} = \frac{\cos^2(\frac{\pi h}{2})}{\sin^2(\frac{\pi h}{2})} = \cot^2\left(\frac{\pi h}{2}\right) = \mathcal{O}(h^{-2}).$$

3 Direct Solvers

3.1 Gaussian Elimination and LU Factorization

The idea of Gaussian elimination: Let us rewrite system (1) $\mathbf{A}\underline{u} = \underline{b}$ in detail as

$$\begin{array}{cccccccccc}
A_{11}^{(0)}u_1 & + & A_{12}^{(0)}u_2 & + \cdots + & A_{1n}^{(0)}u_n & = & b_1^{(0)} \\
A_{21}^{(0)}u_1 & + & A_{22}^{(0)}u_2 & + \cdots + & A_{2n}^{(0)}u_n & = & b_2^{(0)} \\
\vdots & \vdots & \vdots & \vdots \ddots \vdots & \vdots & \vdots & \vdots \\
A_{n1}^{(0)}u_1 & + & A_{n2}^{(0)}u_2 & + \cdots + & A_{nn}^{(0)}u_n & = & b_n^{(0)}.
\end{array}$$

Using the first equation for the elimination of u_1 from the other equations

$$\begin{aligned}
U_{1j} &= A_{1j}^{(0)} = A_{1j}, \ j = 1, 2, \ldots, n, \\
L_{i1} &= A_{i1}^{(0)}/A_{11}^{(0)}, \ i = 2, \ldots, n, \\
A_{ij}^{(1)} &= A_{ij}^{(0)} - L_{i1}U_{1j}, \ i, j = 2, \ldots, n, \\
c_1 &= b_1^{(0)} = b_1, \\
b_i^{(1)} &= b_i^{(0)} - L_{i1}c_1, \ i, j = 2, \ldots, n,
\end{aligned}$$

we arrive at the equivalent system

$$\begin{array}{cccccccc}
U_{11}u_1 & + & U_{12}u_2 & + \cdots + & U_{1n}u_n & = & c_1 \\
 & & A_{22}^{(1)}u_2 & + \cdots + & A_{2N}^{(1)}u_n & = & b_2^{(1)} \\
 & & \vdots & \vdots \ddots \vdots & \vdots & \vdots & \vdots \\
 & & A_{n2}^{(1)}u_2 & + \cdots + & A_{nN}^{(1)}u_n & = & b_n^{(1)}.
\end{array}$$

Now, we can repeat the elimination procedure for the remainder matrix etc. If we simply replace superscript (0) by $(k-1)$ and (1) by (k), then we arrive at the Gaussian (Forward) Elimination Algorithm that transforms the original linear system (1) into an equivalent upper triangular system that can easily be solved by backward substitution.

Algorithm 1 Gaussian Elimination Step

Initialization: $\mathbf{A}^{(0)} = \mathbf{A}$, $\underline{b}^{(0)} = \underline{b}$
Forward Elimination:
for $k = 1, \ldots, n-1$ **do**
  for $j = k, \ldots, n$ **do**
    $U_{kj} = A_{kj}^{(k-1)}$
  end for
  for $i = k+1, \ldots, n$ **do**
    $L_{ik} = A_{ik}^{(k-1)} / A_{kk}^{(k-1)}$
    $b_i^{(k)} = b_i^{(k-1)} - L_{ik} b_k^{(k-1)}$
    for $j = k+1, \ldots, n$ **do**
      $A_{ij}^{(k)} = A_{ij}^{(k-1)} - L_{ik} A_{kj}^{(k-1)}$
    end for
  end for
end for

Storage scheme: The intermediate results after $k-1$ elimination steps can be stored as follows:

$$\begin{pmatrix} U_{11} & U_{12} & \cdots & U_{1k} & \cdots & U_{1n} \\ L_{21} & U_{22} & \cdots & U_{2k} & \cdots & U_{2n} \\ \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ L_{k1} & \cdots & L_{k,k-1} & A_{kk}^{(k-1)} & \cdots & A_{kn}^{(k-1)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ L_{n1} & \cdots & L_{n,k-1} & A_{nk}^{(k-1)} & \cdots & A_{nn}^{(k-1)} \end{pmatrix}$$

Backward substitution: After $n-1$ elimination steps, we obtain the upper triangular system

$$\mathbf{U}\underline{u} = \underline{c}$$

with the upper triangular matrix

$$\mathbf{U} = \begin{pmatrix} U_{11} & U_{12} & \cdots & U_{1,n-1} & U_{1n} \\ 0 & U_{22} & \cdots & U_{2,n-1} & U_{2n} \\ \vdots & 0 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & U_{n-1,n-1} & U_{n-1,n} \\ 0 & 0 & \cdots & 0 & U_{nn} \end{pmatrix} \quad \text{and the rhs } \underline{c} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_{n-1} \\ c_n \end{pmatrix}$$

which can easily be solved by backward substitution:

Algorithm 2 Backward Substitution Step

$u_n = c_n / U_{nn}$
for $i = n-1, n-2, \ldots, 1$ **do**
  $u_i = (c_i - \sum_{j=i+1}^{n} U_{ij} u_j) U_{ii}$
end for

Feasibility: To avoid $U_{kk} = A_{kk}^{k-1} = 0$, i.e., division by zero, we propose a pivot search in the remainder matrix $\mathbf{A}^{(k-1)}$:

1. Total pivoting: column and row exchange defined by
   $i^*, j^* \in \{k, \ldots, n\} : |A_{i^* j^*}^{k-1}| \geq |A_{ij}^{k-1}| \ \ \forall i, j = k, \ldots, n.$
2. Column pivoting: column exchange only.
3. Row pivoting: row exchange only.

Operation count: The SAXPY $(a * x + y)$ operation count yields

1. Forward elimination $\mathbf{A} = \mathbf{LU}$: $\approx \mathcal{O}(n^3) = (n-1)^2 + \cdots + 1^2$,
2. Forward substitution $\underline{c} = \mathbf{L}^{-1}\underline{b}$: $\approx \mathcal{O}(n^2) = (n-1) + \cdots + 1$,
3. Bachward substitution $\underline{x} = \mathbf{U}^{-1}\underline{c}$: $\approx \mathcal{O}(n^2)$.

Gaussian elimination as LU factorization: The $n-1$ Gaussian elimination steps summarized in Algorithm 1 are equivalent to the LU factorization of $\mathbf{A}$, i.e., for instance, in the case $n = 3$, we have

$$\mathbf{A} = \mathbf{LU} = \begin{pmatrix} 1 & 0 & 0 \\ L_{21} & 1 & 0 \\ L_{31} & L_{32} & 1 \end{pmatrix} \begin{pmatrix} U_{11} & U_{21} & U_{31} \\ 0 & U_{22} & U_{32} \\ 0 & 0 & U_{33} \end{pmatrix},$$

with the entries L_{ij} and U_{ij} generated by the Gaussian Elimination Algorithm 1. Therefore, the solution of $\mathbf{A}\underline{u} = \underline{b}$ is equivalent to

1. factorization: $\mathbf{A} = \mathbf{LU}$ by means of $\mathcal{O}(n^3)$ ops,
2. forward substitution: $\mathbf{L}\underline{c} = \underline{b}$ by means of $\mathcal{O}(n^2)$ ops,
3. backward substitution: $\mathbf{U}\underline{u} = \underline{c}$ by means of $\mathcal{O}(n^2)$ ops.

ILU factorization as preconditioner: If we compute the coefficients L_{ij} and U_{ij} in the Gaussian Elimination Algorithm 1 only for the indices

$$(i, j) \in \mathcal{M} \subseteq \mathcal{M}_{all} = \{(i, j) : i, j = 1, 2, \ldots, n\},$$

and set them to zero otherwise, then we obtain an Incomplete LU factorization of the form

$$\mathbf{A} = \widetilde{\mathbf{L}}\widetilde{\mathbf{U}} + \mathbf{R}, \ \text{i.e., in general, } \mathbf{C} = \widetilde{\mathbf{L}}\widetilde{\mathbf{U}} \neq \mathbf{A}.$$

However, $\mathbf{C} = \widetilde{\mathbf{L}}\widetilde{\mathbf{U}}$ can be used as a good *preconditioner* for $\mathbf{A}$ in iterative methods, see Sects. 4 and 5. In practice, the index mask $\mathcal{M}$ is frequently chosen as $\mathcal{M}_{NZE} := \{(i, j) \in \mathcal{M}_{all} : A_{ij} \neq 0\}$, i.e., the LU factorization Algorithm 1 is only performed

on the NonZero Elements. Therefore, there is no fill-in not at all. In particular, $\mathbf{R} = \mathbf{0}$ if $\mathcal{M} = \mathcal{M}_{all}$, and the LU and ILU factorizations coincide.

3.2 Special System Matrices

Symmetric system matrices: The $LDL^\top$ factorization of a symmetric and regular matrix $\mathbf{A}$ can be found by comparing the coefficients ($n = 3$):

$$\begin{aligned}\mathbf{A} &= \begin{pmatrix} 1 & 0 & 0 \\ L_{21} & 1 & 0 \\ L_{31} & L_{32} & 1 \end{pmatrix} \begin{pmatrix} D_{11} & 0 & 0 \\ 0 & D_{22} & 0 \\ 0 & 0 & D_{33} \end{pmatrix} \begin{pmatrix} 1 & L_{21} & L_{31} \\ 0 & 1 & L_{32} \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} D_{11} & D_{11}L_{21} & D_{11}L_{31} \\ L_{21}D_{11} & L_{21}^2 D_{11} + D_{22} & L_{21}L_{31}D_{11} + L_{32}D_{22} \\ L_{31}D_{11} & L_{31}L_{21}D_{11} + L_{32}D_{22} & L_{31}^2 D_{11} + L_{32}^2 D_{22} + D_{33} \end{pmatrix}.\end{aligned}$$

This immediately yields Algorithm 3 providing the $LDL^\top$ factorization in the general case of a symmetric and regular matrix $\mathbf{A}$. We mention that the summation $\sum_{k=1}^{0}$ arising in Algorithm 3 is formally set to 0.

Algorithm 3 $LDL^\top$ factorization

for $j = 1, \ldots, n$ **do**
  $D_{jj} = A_{jj} - \sum_{k=1}^{j-1} L_{jk}^2 D_{kk}$
  for $i = j + 1, \ldots, n$ **do**
    $L_{ij} = D_{jj}^{-1}(A_{jj} - \sum_{k=1}^{j-1} L_{ik} L_{jk} D_{kk})$
  end for
end for

SPD matrices: The Cholesky factorizations $\mathbf{LL}^\top$ or $\mathbf{UU}^\top$ of a SPD matrix $\mathbf{A}$ can also be found by comparing the coefficients ($n = 3$):

$$\begin{aligned}\mathbf{A} &= \begin{pmatrix} L_{11} & 0 & 0 \\ L_{12} & L_{22} & 0 \\ L_{13} & L_{23} & L_{33} \end{pmatrix} \begin{pmatrix} L_{11} & L_{12} & L_{13} \\ 0 & L_{22} & L_{23} \\ 0 & 0 & L_{33} \end{pmatrix} \\ &= \begin{pmatrix} L_{11}^2 & L_{11}L_{12} & L_{11}L_{13} \\ L_{12}L_{11} & L_{12}^2 + L_{22}^2 & L_{12}L_{13} + L_{22}L_{23} \\ L_{13}L_{11} & L_{13}L_{12} + L_{23}L_{22} & L_{13}^2 + L_{23}^2 + L_{33}^2 \end{pmatrix}.\end{aligned}$$

This again yields Algorithm 4 providing the $LL^\top$ factorization in the case when the system matrix $\mathbf{A}$ is SPD. Similarly, one can derive the $UU^\top$ factorization.

Algorithm 4 Cholesky factorizations $\mathbf{L}\mathbf{L}^\top$

$L_{11} = \sqrt{A_{11}}$
for $j = 2, \ldots, n$ **do**
  $L_{1j} = A_{1j}/L_{11}$
  while $j > 2$ **do**
    for $i = 2, \ldots, j-1$ **do**
      $L_{ij} = L_{jj}^{-1}(A_{ij} - \sum_{k=1}^{i-1} L_{ki} L_{kj})$
    end for
  end while
  $L_{jj} = \sqrt{A_{jj} - \sum_{k=1}^{j-1} L_{kj}^2}$
end for

Band and profile matrices: The matrix $\mathbf{A}$ is called band matrix with the bandwidth b_w if $A_{ij} = 0$ for all $|i - j| > b_w$. It easily follows from Algorithm 1 that

$$L_{ij} = 0 \quad \text{and} \quad U_{ij} = 0 \quad \forall\, |i - j| > b_w.$$

Therefore, the following statements are true:

1. The bandwidth of $\mathbf{A}$ remains in the LU factors $\mathbf{L}$ and $\mathbf{U}$ of $\mathbf{A}$, but zero coefficients within the band of $\mathbf{A}$ can turn to non-zero coefficients of $\mathbf{L}$ and $\mathbf{U}$. The latter property is called *"fill-in"*.
2. The factorization needs $\mathcal{O}(b_w^2 n)$ ops, whereas the for- and backward substitutions need only $\mathcal{O}(b_w n)$ ops.
3. The storage requirement is of the order $\mathcal{O}(b_w n)$.

Similar results hold for profile matrices, where we respect different bandwidths from one row to another or from one column to another leading to a row or column *profile* that is sometimes also called *sky line*. As the bandwidth, the row/column respectively column/row profile remains unchanged in the $\mathbf{LU}$ respectively $\mathbf{UL}$ factorization of $\mathbf{A}$.

Sparse direct methods: Sparse direct methods like

- nested dissection methods and
- multifrontal methods

use special elimination strategies that are adapted to the sparsity pattern of the system matrix $\mathbf{A}$ and that can be described as follows:

1. ordering step: reorder the rows and columns,
2. symbolic factorization: nonzero structure of the factors,
3. numerical factorization: $\mathbf{L}$ and $\mathbf{U}$,
4. solution step: forward and backward substitution using $\mathbf{L}$ and $\mathbf{U}$.

In the case of system matrices $\mathbf{K}$ araising from the finite element discretization of boundary value problems like our model problem (2) in $2d$ or $3d$ volumetric computational domains as the unique square or the unique cube, sparse direct methods allow us to reduce the arithmetical complexity to $\mathcal{O}(n^{3/2})$ and $\mathcal{O}(n \log n)$ in $2d$ and

to $\mathcal{O}(n^2)$ and $\mathcal{O}(n^{4/3})$ in $3d$ for the factorization and solution steps, respectively. The memory demand behaves like $\mathcal{O}(n \log n)$ and $\mathcal{O}(n^{4/3})$ in $2d$ and $3d$, respectively.

There are several open-source software packages where different sparse direct solvers are implemented. Let us mention here only the following four packages:

- SuperLU (left-looking): http://crd.lbl.gov/~xiaoye/SuperLU
- UMFPACK (multifrontal): http://www.cise.ufl.edu/research/sparse/umfpack/
- PARDISO (left-right looking): http://www.pardiso-project.org/
- MUMPS (multifrontal): http://mumps.enseeiht.fr/

More information about sparse direct methods can be found, e.g., in the books by George and Liu (1981), Duff et al. (1986), Zlatev (1991) and Davis (2006).

4 Iterative Solvers

4.1 Classical Iteration Methods

General Idea and Questions: Iterative methods obey the following general procedure: Given initial guess $\underline{u}^0 \in \mathbb{R}^n$, generate (how?) successively a sequence of vectors

$$\underline{u}^1, \underline{u}^2, \ldots, \underline{u}^k \longrightarrow \underline{u} \in \mathbb{R}^n : \mathbf{A}\underline{u} = \underline{b} \quad \text{for } k \to \infty\,!$$

In connection with this procedure, the following questions arise:

1. Construction principles?
2. Convergence analysis?
3. Convergence rate and iteration error estimates?
   q-linear: $\exists q \in [0,1)$: $\|\underline{u} - \underline{u}^k\| \le q\|\underline{u} - \underline{u}^{k-1}\| \le q^k\|\underline{u} - \underline{u}^0\|$,
   r-linear: $\exists q \in [0,1)$ and $c = const > 0$: $\|\underline{u} - \underline{u}^k\| \le c\, q^k$.
4. In practice, we use convergence tests, e.g., the defect test

$$\|\underline{d}^k\| = \|\underline{d}^k\|_{\mathbb{R}^n} = \|\underline{e}^k\|_{\mathbf{A}^\top\mathbf{A}} = (\mathbf{A}^\top\mathbf{A}\underline{e}^k, \underline{e}^k)^{0.5}_{\mathbb{R}^n} \le \varepsilon\|\underline{d}^0\|$$

   with the defect $\underline{d}^k = \underline{f} - \mathbf{A}\underline{u}^k = \mathbf{A}(\underline{u} - \underline{u}^k) = \mathbf{A}\underline{e}^k$ and with some $\varepsilon = 10^{-l} \in (0,1)$, to control the iteration. But does the defect test sufficiently well control the Euclidian norm $\|\underline{u} - \underline{u}^k\|_{\mathbb{R}^n}$ of the iteration error $\underline{e}^k = \underline{u} - \underline{u}^k$? The estimate

$$\|\underline{u} - \underline{u}^k\|_{\mathbb{R}^n} = \|\mathbf{A}^{-1}\mathbf{A}(\underline{u} - \underline{u}^k)\|_{\mathbb{R}^n} \le \|\mathbf{A}^{-1}\|_2\,\varepsilon\,\|\mathbf{A}\|_2\,\|\underline{u} - \underline{u}^0\|_{\mathbb{R}^n} \tag{21}$$

   shows that this depends on the spectral condition number $\kappa_2(\mathbf{A})$ of the system matrix $\mathbf{A}$.
5. What is the right choice of the norm $\|\cdot\|$ in which we control the error $\underline{e}^k$ of the iteration procedure?

The Jacobi iteration: If we resolve the ith equation $A_{i1}u_1 + \cdots + A_{ii}u_i + \cdots + A_{in}u_n = b_i$ of (1) for u_i, then we get the fixed point equation $u_i = A_{ii}^{-1}(f_i - \sum_{j\neq i} A_{ij}u_j)$ immediately yielding the following fixed point iteration given in Algorithm 5. As our analysis will show, in the case of finite element equations ($\mathbf{A} = \mathbf{K}$), the Jacobi method exhibits slow convergence, but the damped version has excellent smoothing properties. Therefore, the damped Jacobi method is often used as smoother in multigrid methods, see also Sect. 6. The Jacobi methods can be seen as prototype of Additive Schwarz Methods, see Sect. 5.

Algorithm 5 Jacobi iteration method

Given initia guess $\underline{u}^0 = (u_1^0, \ldots, u_n^0)^\top \in \mathbb{R}^n$
for $k = 0, \ldots, k_{stop}$ until convergence (defect test) **do**
$\underline{u}^{k+1} = (u_1^{k+1}, \ldots, u_n^{k+1})^\top \in \mathbb{R}^n$:

$$u_i^{k+1} = \frac{1}{A_{ii}}\left(f_i - \sum_{j=1, j\neq i}^{n} A_{ij}u_j^k\right) \quad \text{for } i = 1, 2, \ldots, n \text{ (in parallel)}$$

end for

The Gauss–Seidel iteration: If we use the already computed new components $u_1^{k+1}, \ldots, u_{i-1}^{k+1}$ in the Jacobi iteration, then we arrive at the following iterative procedure known as Gauss–Seidel iteration, see Algorithm 6.

Algorithm 6 Gauss–Seidel iteration method

Given initia guess $\underline{u}^0 = (u_1^0, \ldots, u_n^0)^\top \in \mathbb{R}^n$
for $k = 0, \ldots, k_{stop}$ until convergence (defect test) **do**
$\underline{u}^{k+1} = (u_1^{k+1}, \ldots, u_n^{k+1})^\top \in \mathbb{R}^n$:

$$u_i^{k+1} = \frac{1}{A_{ii}}\left(f_i - \sum_{j=1}^{i-1} A_{ij}u_j^{k+1} - \sum_{j=i+1}^{n} A_{ij}u_j^k\right)$$

for $i = 1, 2, \ldots, n$ (sequentially)
end for

As the Jacobi method, in the case of finite element equations ($\mathbf{A} = \mathbf{K}$), the Gauss–Seidel method also exhibits slow convergence, but it has excellent smoothing properties. Therefore, the Gauss–Seidel iteration is most frequently used as smoother in multigrid methods, see also Sect. 6. The Gauss–Seidel method can be seen as prototype of Multiplicative Schwarz Methods, see Sect. 5.

Richardson and preconditioned Richardson iteration methods: Let us give the following motivation. Solving the ODE system

$$\frac{\partial \underline{u}(t)}{\partial t} + \mathbf{A}\underline{u}(t) = \underline{b}$$

by the explicit Euler method gives the *Richardson method*

$$\frac{\underline{u}^{k+1} - \underline{u}^k}{\tau} + \mathbf{A}\underline{u}^k = \underline{b}, \quad k = 1, 2, \ldots. \tag{22}$$

Application of Richardson (22) to the preconditioned (reduced stiffness) system

$$\mathbf{C}^{-1}\mathbf{A}\underline{u} = \mathbf{C}^{-1}\underline{b} \quad \Longleftrightarrow \quad \mathbf{A}\underline{u} = \underline{b}$$

gives the *preconditioned Richardson method*

$$\mathbf{C}\frac{\underline{u}^{k+1} - \underline{u}^k}{\tau} + \mathbf{A}\underline{u}^k = \underline{b}, \quad k = 1, 2, \ldots \tag{23}$$

where the preconditioner $\mathbf{C}$ should reduce the stiffness and should be easily invertible at the same time, see also Sect. 5. The convergence rate heavily depends on the quality of the preconditioner, see the analysis given below!

Algorithm 7 Preconditioned Richardson method

Given initia guess $\underline{u}^0 = (u_1^0, \ldots, u_n^0)^\top \in \mathbb{R}^n$
for $k = 0, \ldots, k_{stop}$ until convergence (defect test) **do**

$$\begin{aligned} \underline{d}^k &= \underline{f} - \mathbf{A}\underline{u}^k \\ \mathbf{C}\underline{w}^k &= \underline{d}^k \\ \underline{u}^{k+1} &= \underline{u}^k + \tau \underline{w}^k \end{aligned}$$

end for

Special choices of the preconditioner $\mathbf{C}$ in the preconditioned Richardson method (23) yield well-known classical iteration methods:

1. $\mathbf{C} = \mathbf{I}$: Classical Richardson method,
2. $\mathbf{C} = \mathbf{D} := \operatorname{diag} \mathbf{A}$: τ-Jacobi method ($\tau = 1$: Jacobi method),
3. $\mathbf{C} = \mathbf{L} + (1/\omega)\mathbf{D}$: SOR preconditioner ($\mathbf{A} = \mathbf{L} + \mathbf{D} + \mathbf{U}$):
   $\tau = 1$: SOR = Successive OverRelaxation (D. Young 1950),
   $\tau = 1$ and $\omega = 1$: Gauss–Seidel,
4. $\mathbf{C} = \widetilde{\mathbf{L}}\widetilde{\mathbf{U}}$: ILU decomposition of $\mathbf{A}$, see Sect. 3,
5. Modern preconditioners, see Sect. 5.

From the preconditioned Richardson iteration (23) and $\mathbf{A}\underline{u} = \underline{b}$, we can immediately derive the error iteration scheme

$$\underline{e}^{k+1} = \underline{u} - \underline{u}^{k+1} = \underline{u} - (\underline{u}^k - \tau \mathbf{C}^{-1}\mathbf{A}(\underline{u} - \underline{u}^k)) = \mathbf{E}\,\underline{e}^k \tag{24}$$

with the error propagation (iteration) matrix $\mathbf{E} := \mathbf{I} - \tau\mathbf{C}^{-1}\mathbf{A}$. The error iteration scheme (24) has the following consequences wrt convergence:

1. The Richardson iteration (22) converges iff the spectral radius $\rho(\mathbf{E}) := \max_{i=1,\dots,n} |\lambda_i(\mathbf{E})|$ of $\mathbf{E}$ is less than 1.
2. Error estimate wrt some norm and q-linear convergence:

$$\|\underline{e}^{k+1}\| \le \|\mathbf{I} - \tau\mathbf{C}^{-1}\mathbf{A}\|\,\|\underline{e}^k\| = q\,\|\underline{e}^k\| \le q^{k+1}\|\underline{e}^0\| \to 0$$

provided that $q = \|\mathbf{E}\| < 1$ in some norm $\|\cdot\|$!

Let us now consider the SPD case, i.e., let us assume that $\mathbf{A}$ and $\mathbf{C}$ are SPD. Then, we have

$$\mathbf{A}\underline{u} = \underline{f} \quad \Longleftrightarrow \quad \widetilde{\mathbf{A}}\underline{\tilde{u}} = \underline{\tilde{f}}$$

with $\underline{\tilde{f}} = \mathbf{C}^{-1/2}\underline{f}$, $\underline{\tilde{u}} = \mathbf{C}^{1/2}\underline{u}$, and the preconditioned stiffness matrix $\widetilde{\mathbf{A}} = \mathbf{C}^{-1/2}\mathbf{A}\mathbf{C}^{-1/2}$ that is obviously SPD. Thus it is sufficient to derive iteration error estimates for the classical Richardson method (22). Let us consider expansion of the kth error $\underline{e}^k$ into a Fourier series wrt the eigenvectors of $\mathbf{A}$ (resp. $\widetilde{\mathbf{A}}$):

$$\underline{e}^k = \sum_{j=1}^{n} \alpha_j \underline{\varphi}_j \tag{25}$$

with the Fourier coefficients $\alpha_j = (\underline{e}^k, \underline{\varphi}_j)_{\mathbb{R}^n}$, $j = 1, 2, \dots, n$. Inserting the Fourier expansion (25) into the error scheme (24) with $\mathbf{C} = \mathbf{I}$, we get

$$\underline{e}^{k+1} = \mathbf{E}\underline{e}^k = (\mathbf{I} - \tau\mathbf{A})\underline{e}^k = \sum_{j=1}^{n} \alpha_j (1 - \tau\lambda_j)\underline{\varphi}_j. \tag{26}$$

We now choose the following class of norms

$$\|\underline{v}\|_s = \|\underline{v}\|_{\mathbf{A}^s} := (\mathbf{A}^s\underline{v}, \underline{v})^{1/2}_{\mathbb{R}^n}, \quad s \in \mathbb{R} \quad \text{(special interest: } s = 0, 1, 2),$$

in which we will derive sharp iteration error estimates. Using (26), we get the sharp estimate

$$
\begin{aligned}
\|\underline{e}^{k+1}\|_s^2 &= (\mathbf{A}^s \underline{e}^{k+1}, \underline{e}^{k+1})_{\mathbb{R}^n} = (\mathbf{A}^s \underline{e}^{k+1}, \underline{e}^{k+1}) \\
&= \left(\mathbf{A}^s \sum_{j=1}^{n} \alpha_j (1-\tau\lambda_j) \underline{\varphi}_j, \sum_{i=1}^{n} \alpha_i (1-\tau\lambda_i) \underline{\varphi}_i \right) \\
&= \left(\sum_{j=1}^{n} \alpha_j (1-\tau\lambda_j) \lambda_j^s \underline{\varphi}_j, \sum_{i=1}^{n} \alpha_i (1-\tau\lambda_i) \underline{\varphi}_i \right) \\
&= \sum_{j=1}^{n} \alpha_j^2 \lambda_j^s (1-\tau\lambda_j)^2 \\
&\leq \max_{i=1,\ldots,n} (1-\tau\lambda_i)^2 \sum_{j=1}^{n} \alpha_j^2 \lambda_j^s = \max_{i=1,\ldots,n} (1-\tau\lambda_i)^2 \, \|\underline{e}^k\|_s^2 \\
&= (\max\{|1-\tau\lambda_1|, |1-\tau\lambda_n|\})^2 \|\underline{e}^k\|_s^2 = q(\tau)^2 \|\underline{e}^k\|_s^2.
\end{aligned}
$$

Lemma 4.1 (Convergence rate estimate) *The* $\|\cdot\|_s$ *norm of the iteration matrix* $\mathbf{E} = \mathbf{I} - \tau\mathbf{A}$ *is given by*

$$
\|\mathbf{E}\|_s := \max_{\underline{0} \neq \underline{v} \in \mathbb{R}^n} \frac{\|\mathbf{E}\underline{v}\|_s}{\|\underline{v}\|_s} = q(\tau) := \max\{|1-\tau\lambda_1|, |1-\tau\lambda_n|\} < 1
$$

for fixed $\tau \in (0, 2/\lambda_n)$ *and* $s \in \mathbb{R}$.

Remark 4.2 $\|\underline{e}^{k+1}\|_s \leq q(\tau)\|\underline{e}^k\|_s \leq \cdots \leq (q(\tau))^{k+1}\|\underline{e}^0\|_s$
$s = 0 : \|\underline{u} - \underline{u}^{k+1}\|_{\mathbb{R}^n} \leq (q(\tau))^{k+1}\|\underline{u} - \underline{u}^0\|_{\mathbb{R}^n}$ (not computable!)
$s = 1 : \|\underline{u} - \underline{u}^{k+1}\|_{\mathbf{A}} \leq (q(\tau))^{k+1}\|\underline{u} - \underline{u}^0\|_{\mathbf{A}}$ (not computable!)
$s = 2 : \|\underline{u} - \underline{u}^{k+1}\|_{\mathbf{A}^2} = \|\underline{d}^{k+1}\|_{\mathbb{R}^n} \leq (q(\tau))^{k+1}\|\underline{d}^0\|_{\mathbb{R}^n}$ (computable!)

We immediately see from Fig. 1 that the optimal iteration parameter $\tau_{opt} = 2/(\lambda_1 + \lambda_n)$ follows from the equation $1 - \tau\lambda_1 = \tau\lambda_n - 1$. Therefore, the optimal rate is given by the formula

$$
q_{opt} = q(\tau_{opt}) = \frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1} = \frac{\kappa_2 - 1}{\kappa_2 + 1} \tag{27}
$$

with the spectral condition number $\kappa_2 = \kappa_2(\mathbf{A}) = \lambda_n/\lambda_1 = \frac{\lambda_{max}(\mathbf{A})}{\lambda_{min}(\mathbf{A})}$.

Theorem 4.3 (Optimal convergence rates) *In the SPD case, the classical Richardson method (22) converges for all* $\tau \in (0, 2/\lambda_{max}(\mathbf{A})) = (0, 2/\lambda_n)$*, and, for every fixed* $s \in \mathbb{R}$*, the iteration error estimate*

$$
\|\underline{u} - \underline{u}^{k+1}\|_s \leq q(\tau)\|\underline{u} - \underline{u}^k\|_s
$$

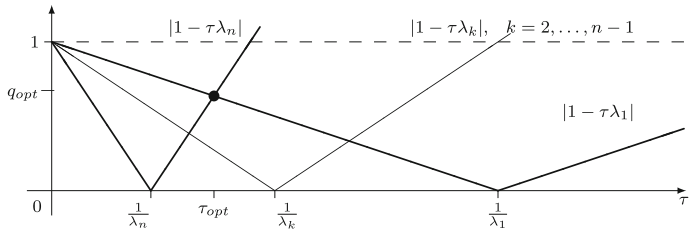


Fig. 1 Functions $|1-\tau\lambda_k|$

holds with $q(\tau) := \max\{|1-\tau\lambda_1|, |1-\tau\lambda_n|\} < 1$. *The optimal (minimal) rate*

$$q_{opt} = q(\tau_{opt}) = \frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1} = \frac{\kappa_2(\mathbf{A}) - 1}{\kappa_2(\mathbf{A}) + 1}$$

is attained at $\tau_{opt} = 2/(\lambda_1 + \lambda_n)$, *where* λ_1 *and* λ_n *are the minimal and maximal eigenvalues of the matrix* $\mathbf{A}$, *respectively.*

Since the preconditioned Richardson method (23) can be interpreted as the application of the classical Richardson method (22) to the preconditioned system

$$\widetilde{\mathbf{A}}\underline{\tilde{u}} = \underline{\tilde{b}} \quad \Longleftrightarrow \quad \mathbf{A}\underline{u} = \underline{b}$$

with $\widetilde{\mathbf{A}} = \mathbf{C}^{-1/2}\mathbf{K}\mathbf{C}^{-1/2}$, $\underline{\tilde{b}} = \mathbf{C}^{-1/2}\underline{b}$, $\underline{\tilde{u}} = \mathbf{C}^{1/2}\underline{u}$, we get the same convergence results as presented in the Theorem 4.3, but now with

$$\lambda_1 = \lambda_{min}(\mathbf{C}^{-1/2}\mathbf{A}\mathbf{C}^{-1/2}) = \lambda_{min}(\mathbf{C}^{-1}\mathbf{A}),$$

$$\lambda_n = \lambda_{max}(\mathbf{C}^{-1/2}\mathbf{A}\mathbf{C}^{-1/2}) = \lambda_{max}(\mathbf{C}^{-1}\mathbf{A}).$$

Remark 4.4 The choice $s = 2$ gives a norm in which the norm of the kth iteration error $\underline{\tilde{e}}^k$ is computable. Indeed,

$$\|\underline{\tilde{u}} - \underline{\tilde{u}}^k\|_2^2 = (\widetilde{\mathbf{A}}^2\underline{\tilde{e}}^k, \underline{\tilde{e}}^k) = (\mathbf{A}\mathbf{C}^{-1}\mathbf{A}\underline{e}^k, \underline{e}^k) = (\underline{w}^k, \underline{d}^k).$$

If $\mathbf{C}$ is close to $\mathbf{A}$, then this norm is obviously close to the $\mathbf{A}$-energy norm, in which one often wants to control the iteration error.

Remark 4.5 If we solve the finite element equation (12), with the stiffness matrix $\mathbf{K}$ as system matrix $\mathbf{A}$, by means of the τ-Jacobi that is nothing but the preconditioned Richardson method (23) with $\mathbf{C} = \mathbf{D} := \operatorname{diag}\mathbf{K}$, we get the optimal rate

$$q_{opt} = q(\tau_{opt}) = \frac{\kappa_2(\mathbf{C}^{-1}\mathbf{K}) - 1}{\kappa_2(\mathbf{C}^{-1}\mathbf{K}) + 1} = \frac{1 - \mathcal{O}(h^2)}{1 + \mathcal{O}(h^2)} = 1 - \mathcal{O}(h^2)$$

at the optimal iteration parameter τ_{opt} since $\kappa_2(\mathbf{C}^{-1}\mathbf{K}) = \mathcal{O}(h^{-2})$ due to estimate (18).

Example 4.6 Let us consider Example 2.1. Then we immediately see the following statements:

1. $\mathbf{A} = \mathbf{K} = h^{-1}\,\text{tridiag}\,(-1, 2, -1)$,
2. $\mathbf{C} = \text{diag}\,\mathbf{K} = \text{diag}\,\left(2h^{-1}\right)$,
3. $\lambda_{min}(\mathbf{C}^{-1}\mathbf{K}) = (h/2)\,4h^{-1}\sin^2(\frac{\pi h}{2}) = 2\sin^2(\frac{\pi h}{2})$
4. $\lambda_{max}(\mathbf{C}^{-1}\mathbf{K}) = (h/2)\,4h^{-1}\cos^2(\frac{\pi h}{2}) = 2\cos^2(\frac{\pi h}{2})$,
5. $\tau_{opt} = \frac{2}{\lambda_{min}+\lambda_{max}} = \frac{2}{2(\sin^2(\pi h/2)+\cos^2(\pi h/2))} = 1$,
6. the classical Jacobi iteration is optimal wrt τ!
7. $q_{opt} = \frac{\kappa_2(\mathbf{C}^{-1}\mathbf{K})-1}{\kappa_2(\mathbf{C}^{-1}\mathbf{K})+1} = \frac{1-\tan^2(\frac{\pi h}{2})}{1+\tan^2(\frac{\pi h}{2})} = 1 - 2\sin^2(\frac{\pi h}{2}) \approx 1 - \frac{\pi^2 h^2}{2}$.

Therefore, the Jacobi method converges very slowly. More precisely, we need $I(\varepsilon) = \mathcal{O}(h^{-2}\ln\varepsilon^{-1})$ iterations in order to reduce the initial error by the factor $\varepsilon \in (0, 1)$. However, the damped Jacobi method, e.g., $\tau = 1/\lambda_{max}(\mathbf{C}^{-1}\mathbf{K}) = 1/\lambda_n = 1/(2\cos^2(\pi/2)) \approx 1/2$, leads to a fast reduction of the high frequency modes, see mode reduction rates $|1 - \tau\lambda_k|$ in Fig. 1. This property turns the damped Jacobi method into a perfect smoother in multigrid methods, see Sect. 6.

In practice, we use

- preconditioned Krylov subspace iteration methods

instead of preconditioned Richardson iteration methods since

1. they don't need spectral information to determine iteration parameters like τ in Richardson, and
2. they converge faster!

In the SPD case,

- Preconditioned Conjugate Gradient (PCG) method,

is the method of choice. The Conjugate Gradient (CG) methods was proposed by Hestenes and Stiefel (1952), and is now one of the most popular methods in numerical linear algebra (number 3 under the top 10 numerical algorithms[1]).

4.2 Gradient and Conjugate Gradient Methods for SPD Systems

SPD systems and minimization problems: Let us consider the linear system: Find $\underline{u} \in \mathbb{R}^n$ such that

$$\mathbf{A}\underline{u} = \underline{b} \tag{28}$$

[1] http://www.uta.edu/faculty/rcli/TopTen/topten.pdf.

with given rhs $\underline{b} \in \mathbb{R}^n$ and SPD system matrix $\mathbf{A}$, i.e., $\mathbf{A} = \mathbf{A}^\top$ and $(\mathbf{A}\underline{v}, \underline{v}) > 0 \; \forall \underline{v} \in \mathbb{R}^n : \underline{v} \neq \underline{0}$. Then the SPD system (28) is equivalent to the energy minimization problem

$$J(\underline{u}) = \min_{\underline{v} \in \mathbb{R}^n} \frac{1}{2}(\mathbf{A}\underline{v}, \underline{v}) - (\underline{b}, \underline{v}) = \min_{\underline{v} \in \mathbb{R}^n} \frac{1}{2} \sum_{i,j=1}^{n} A_{ij} v_j v_i - \sum_{i=1}^{n} b_i v_i$$

that is nothing but the minimization problem for a quadratic function of n variables $v_1, \ldots, v_n$. Indeed, we have

$$\nabla J(\underline{v}) = \left(\frac{\partial J(\underline{v})}{\partial v_i} \right)_{i=1,\ldots,n} = \mathbf{A}\underline{v} - \underline{b} = \underline{0}$$

and

$$\nabla^2 J(\underline{v}) = \left(\frac{\partial^2 J(\underline{v})}{\partial v_i \partial v_j} \right)_{i,j=1,\ldots,n} = \mathbf{A} \quad \text{is SPD,}$$

that are nothing but the necessary and sufficient condition for a minimum.

Gradient (steepest descent) method: The idea for a Gradient (Steepest Descent) Method can be summarized as follows:

1. Given initial guess $\underline{u}^0 = (u_1^0, \ldots, u_n^0)^\top \in \mathbb{R}^n$,
2. compute steepest descent $\underline{d}^0$ at $\underline{u}^0$: $\underline{d}^0 = -\nabla J(\underline{u}^0) = \underline{b} - \mathbf{A}\underline{u}^0$,
3. $\underline{s}^0 = \underline{d}^0$ (search direction),
4. $\underline{u}^1 = \underline{u}^0 + \alpha_1 \underline{s}^0$ (next iterate),
5. compute the step size α_1 such that $J(\underline{u}^0 + \alpha_1 \underline{s}^0) = \min_\alpha J(\underline{u}^0 + \alpha \underline{s}^0)$, i.e., $\frac{dJ(\underline{u}^0 + \alpha \underline{s}^0)}{d\alpha} = (\mathbf{A}\underline{u}^0, \underline{s}^0) - (\underline{f}, \underline{s}^0) + \alpha\,(\mathbf{A}\underline{s}^0, \underline{s}^0) = 0$ gives

$$\alpha_1 = (\underline{d}^0, \underline{s}^0)/(\mathbf{A}\underline{s}^0, \underline{s}^0).$$

The new steepest descent $\underline{d}^1$ at $\underline{u}^1$ can be computed by recursion as follows

$$\underline{d}^1 = \underline{b} - \mathbf{A}\underline{u}^1 = \underline{b} - \mathbf{A}(\underline{u}^0 + \alpha_1 \underline{s}^0) = \underline{b} - \mathbf{A}\underline{u}^0 - \alpha_1 \mathbf{A}\underline{s}^0 = \underline{d}^0 - \alpha_1 \mathbf{A}\underline{s}^0.$$

Using this recursion, we arrive at Algorithm 8.

Since

$$\begin{aligned} J(\underline{v}) &= 0.5(\mathbf{A}\underline{v}, \underline{v}) - (\underline{f}, \underline{v}) \\ &= 0.5(\mathbf{A}\underline{v}, \underline{v}) - (\mathbf{A}\underline{u}, \underline{v}) + 0.5(\mathbf{A}\underline{u}, \underline{u}) - 0.5(\mathbf{A}\underline{u}, \underline{u}) \\ &= 0.5\|\underline{u} - \underline{v}\|_{\mathbf{A}}^2 - 0.5\|\underline{u}\|_{\mathbf{A}}^2, \end{aligned}$$

we conclude that

$$\min_{\underline{v} \in \mathbb{R}^n} J(\underline{v}) \quad \Leftrightarrow \quad \min_{\underline{v} \in \mathbb{R}^n} \|\underline{u} - \underline{v}\|_{\mathbf{A}},$$

Algorithm 8 Gradient method = steepest descent method

Initialization:
$\underline{u}^0 = (u_1^0, \ldots, u_n^0)^\top \in \mathbb{R}^n$ - given initial guess
$\underline{d}^0 = \underline{b} - \mathbf{A}\underline{u}^0$ - initial defect = steepest descent
$\underline{s}^0 = \underline{d}^0$ - search direction
Iteration:
for $k = 0, \ldots, k_{stop}$ **do**
  if $\|\underline{d}^k\| \le \varepsilon \|\underline{d}^0\|$ **then**
    STOP (defect test)
  else
    $\alpha_k = (\underline{d}^k, \underline{s}^k)/(\mathbf{A}\underline{s}^k, \underline{s}^k)$ - new step size
    $\underline{u}^{k+1} = \underline{u}^k + \alpha_k \underline{s}^k$ - new iterate
    $\underline{d}^{k+1} = \underline{d}^k - \alpha_k \mathbf{A}\underline{s}^k$ - new defect
    $\underline{s}^{k+1} = \underline{d}^{k+1}$ - new search direction
  end if
end for

where $\|\underline{u}\|_{\mathbf{A}} := (\mathbf{A}\underline{u}, \underline{u})^{1/2}$ denotes the energy norm. Using $\underline{s}^k = \underline{d}^k$, we get

$$\begin{aligned}\|\underline{u} - \underline{u}^{k+1}\|_{\mathbf{A}} &= \min_{\alpha \in \mathbb{R}} \|\underline{u} - (\underline{u}^k + \alpha \underline{s}^k)\|_{\mathbf{A}} \\ &\le \|\underline{u} - (\underline{u}^k + \tau_{opt}\underline{d}^k)\|_{\mathbf{A}} \le q(\tau_{opt})\|\underline{u} - \underline{u}^k\|_{\mathbf{A}}.\end{aligned}$$

Thus, the gradient method converges at least as fast as the Richardson method.

The two following improvements of the gradient method are possible:

1. Preconditioning: Apply the gradient method to the preconditioned system

$$\mathbf{C}^{-1}\mathbf{A}\underline{u} = \mathbf{C}^{-1}\underline{b} \iff \mathbf{C}^{-0.5}\mathbf{A}\mathbf{C}^{-0.5}\underline{v} = \mathbf{C}^{-0.5}\underline{b}.$$

   This means that the search direction in the Preconditioned Gradient Method is the preconditioned defect

$$\underline{s}^{k+1} = \underline{w}^{k+1} := \mathbf{C}^{-1}\underline{d}^{k+1}.$$

2. Use conjugate search directions defined by

$$\underline{s}^{k+1} = \underline{d}^{k+1} + \beta_k \underline{s}^k \ \perp \ \underline{s}^k \text{ wrt } (\cdot,\cdot)_{\mathbf{A}} := (\mathbf{A}\cdot, \cdot),$$

   i.e., $\beta_k \in \mathbb{R} : (\mathbf{A}\underline{s}^{k+1}, \underline{s}^k) = 0 \Rightarrow \beta_k = -(\mathbf{A}\underline{d}^{k+1}, \underline{s}^k)/(\mathbf{A}\underline{s}^k, \underline{s}^k)$.

Preconditioned Conjugate Gradient Method: Both improvements lead to the Preconditioned Conjugate Gradient (PCG) Method presented in Algorithm 9.

Algorithm 9 Preconditioned conjugate gradient method

Initialization:
$\underline{u}^0 = (u_1^0, \dots, u_n^0)^\top \in \mathbb{R}^n$ - given initial guess
$\underline{d}^0 = \underline{b} - \mathbf{A}\underline{u}^0$ - initial defect = steepest descent
$\underline{s}^0 = \underline{w}^0 := \mathbf{C}^{-1}\underline{d}^0$ - search direction = preconditioned defect
Iteration:
for $k = 0, \dots, k_{stop}$ **do**
  if $\|\underline{e}^k\|_{\mathbf{AC}^{-1}\mathbf{A}} \le \varepsilon \|\underline{e}^0\|_{\mathbf{AC}^{-1}\mathbf{A}}$ **then**
    STOP ($\mathbf{AC}^{-1}\mathbf{A}$ - norm test)
  else
    $\alpha_k = (\underline{d}^k, \underline{s}^k)/(\mathbf{A}\underline{s}^k, \underline{s}^k) = (\underline{d}^k, \underline{w}^k)/(\mathbf{A}\underline{s}^k, \underline{s}^k)$ - new step size
    $\underline{u}^{k+1} = \underline{u}^k + \alpha_k \underline{s}^k$ - new iterate
    $\underline{d}^{k+1} = \underline{d}^k - \alpha_k \mathbf{A}\underline{s}^k$ - new defect
    $\underline{w}^{k+1} := \mathbf{C}^{-1}\underline{d}^{k+1}$ - preconditioning
    $\beta_k = -(\mathbf{A}\underline{w}^{k+1}, \underline{s}^k)/(\mathbf{A}\underline{s}^k, \underline{s}^k) = (\underline{w}^{k+1}, \underline{d}^{k+1})/(\underline{w}^k, \underline{d}^k)$
    $\underline{s}^{k+1} = \underline{w}^{k+1} + \beta_k \underline{s}^k$ - new search direction
  end if
end for

The $\mathbf{AC}^{-1}\mathbf{A}$ norm test $\|\underline{e}^k\|_{\mathbf{AC}^{-1}\mathbf{A}} \le \varepsilon \|\underline{e}^0\|_{\mathbf{AC}^{-1}\mathbf{A}}$ is nothing but

$$(\underline{w}^k, \underline{d}^k) \le \varepsilon (\underline{w}^0, \underline{d}^0), \tag{29}$$

and can easily be computed by the Euclidian product of the preconditioned defect $\underline{w}^k$ and the defect $\underline{d}^k$. In the case of a good and appropriately scaled preconditioner $\mathbf{C}$, we can assume that $\mathbf{C}^{-1}\mathbf{A} \approx \mathbf{I}$, and, therefore, the computable test in the $\mathbf{AC}^{-1}\mathbf{A}$ norm is very close to the test $\|\underline{e}^k\|_{\mathbf{A}} \le \varepsilon \|\underline{e}^0\|_{\mathbf{A}}$ wrt the $\mathbf{A}$-energy norm that is often the norm in which we want to control the error.

Theorem 4.7 (PCG: convergence rate estimate) *Let* $\mathbf{A}$ *and* $\mathbf{C}$ *be SPD matrices. Then not more than*

$$I(\varepsilon) = [|\ln(\varepsilon^{-1} + (\varepsilon^{-2} + 1)^{0.5})/\ln(\tilde{q}^{-1})|]$$

iteration are necessary to reduce the initial error $\|\underline{u} - \underline{u}^0\|_{\mathbf{A}}$ *by the factor* $\varepsilon \in (0, 1)$. *Moreover, the iteration error estimate*

$$\|\underline{u} - \underline{u}^{k+1}\|_{\mathbf{A}} \le \eta^{(k+1)} \|\underline{u} - \underline{u}^0\|_{\mathbf{A}} \tag{30}$$

holds, where

$$\eta^{(k+1)} := \frac{2q^{k+1}}{1 + q^{2(k+1)}} \le 2q^{k+1}, \quad \text{with } q = \frac{\sqrt{\kappa_2(\mathbf{C}^{-0.5}\mathbf{AC}^{-0.5})} - 1}{\sqrt{\kappa_2(\mathbf{C}^{-0.5}\mathbf{AC}^{-0.5})} + 1} < 1.$$

Proof It is enough to investigate the convergence of the PCG method for the unpreconditioned case $\mathbf{C} = \mathbf{I}$ since the PCG is nothing but the CG applied to the preconditioned system. The $k+1$ CG iterate $\underline{u}^{k+1}$ minimizes the energy functional $J(\underline{v})$

over the set $\underline{u}^0 + \mathcal{K}_{k+2}(\mathbf{A}, \underline{d}^0)$, where

$$\begin{aligned}\mathcal{K}_{k+2}(\mathbf{A}, \underline{d}^0) &= \text{span}\{\underline{d}^0, \mathbf{A}\underline{d}^0, \ldots, \mathbf{A}^{k+1}\underline{d}^0\} \\ &= \text{span}\{\underline{d}^0, \underline{d}^1, \ldots, \underline{d}^{k+1}\}\end{aligned}$$

is the so-called Krylov subspace. Therefore, we have

$$\begin{aligned}\|\underline{u} - \underline{u}^{k+1}\|_{\mathbf{A}} &= \min_{\underline{v} \in \underline{u}^0 + \mathcal{K}_{k+2}(\mathbf{A}, \underline{d}^0)} \|\underline{u} - \underline{v}\|_{\mathbf{A}} \\ &\leq \min_{p_{k+1} \in \mathcal{P}_{k+1}:\, p_{k+1}(0)=1} \|p_{k+1}(\mathbf{A})\underline{e}^0\|_{\mathbf{A}} \\ &\leq \left[\min_{p_{k+1} \in \mathcal{P}_{k+1}:\, p_{k+1}(0)=1} \max_{i=1,\ldots,n} |p_{k+1}(\lambda_i)| \right] \|\underline{e}^0\|_{\mathbf{A}},\end{aligned}$$

where [...] is (almost) the famous Chebychev approximation problem, i.e., PCG converges at least as fast as the Chebychev method. This observation immediately leads to estimate (30). We refer the reader to Saad (2003) for a more detailed presentation of the proof. □

5 Preconditioners

5.1 *Basic Idea of Preconditioning*

As already mentioned in the previouse sections, the basic idea of preconditioning can be viewed in the following way: We multiply the linear system

$$\mathbf{A}\underline{u} = \underline{b},$$

with the inverse of a regular matrix $\mathbf{C}$, i.e.,

$$\mathbf{C}^{-1}\mathbf{A}\underline{u} = \mathbf{C}^{-1}\underline{b}. \tag{31}$$

Then the application of standard iterative schemes like the Richardson method or the Conjugate Gradient method lead to the preconditioned versions, see Algorithm 7 and Algorithm 9. Since the convergence of these iterative schemes heavily depends on the condition number $\kappa_2(\mathbf{C}^{-1}\mathbf{A})$, an efficient preconditioner should fulfill the following requirements:

- Reduce the condition number $\kappa_2(\mathbf{C}^{-1}\mathbf{A}) \ll \kappa_2(\mathbf{A})$ as much as possible, if feasible, then $\kappa_2(\mathbf{C}^{-1}\mathbf{A})$ should be independent of h.
- Cheap realization of the operation $\mathbf{C}^{-1}\underline{d}$, i.e., with complexity of

$$\mathcal{O}(n_h) \quad \text{or} \quad \mathcal{O}(n_h \log^{\alpha}(n_h)) \quad \text{arithmetical operations.}$$

To obtain bounds for the condition number, we can use the following Lemma, which can be proved by using the Rayleigh quotient (compare also estimate (18)).

Lemma 5.1 *For* $\mathbf{A}, \mathbf{C} \in \mathbb{R}^{n_h \times n_h}$ *symmetric and positive definite, let the spectral equivalence inequalities*

$$\mu_1(\mathbf{C}\underline{v}, \underline{v}) \leq (\mathbf{A}\underline{v}, \underline{v}) \leq \mu_2(\mathbf{C}\underline{v}, \underline{v}) \quad \forall \underline{v} \in \mathbb{R}^{n_h}$$

be fulfilled with some positive constants μ_1 *and* μ_2*. Then there holds the estimate*

$$\kappa_2(\mathbf{C}^{-1}\mathbf{A}) \leq \frac{\mu_2}{\mu_1}.$$

There are usually two classes of preconditioners, i.e., one class that uses only information about the system matrix $\mathbf{A}$, and preconditioners which use information coming from the underlying variational problem (see also the heredity relation (13)):

- Algebraic preconditioners:
  - Incomplete LU-factorization (ILU)
  - Incomplete Cholesky-factorization (IC)
  - Algebraic multigrid method (AMG)
  - ...
- Preconditioners using variational background:
  - Schwarz methods
  - Multilevel methods (BPX, MDS, AMLI,...)
  - Multigrid methods (GMG, AMG)
  - Domain decomposition methods (DDM)
  - ...

In the next subsection, we will focus on preconditioners with variational background.

5.2 Subspace Correction Methods

Here we will use the fact that the linear system which we want to solve is equivalent to a discrete variational problem (see also Sect. 2), i.e.,

$$\mathbf{K}\underline{u} = \underline{f} \quad \Leftrightarrow \quad u^h \in V_0^h : a(u^h, v^h) = \ell(v^h) \; \forall v^h \in V_0^h,$$

with the representation of a function by the coefficient vector

$$u^h = \sum_{j=1}^{n_h} u_j N_j(\mathbf{x}) \in V_0^h \quad \leftrightarrow \quad \underline{u} \in \mathbb{R}^{n_h}.$$

Note that we also consider the problem already in homogenized form.

A first idea: To obtain the spectral equivalence estimate of Lemma 5.1 for some preconditioner $\mathbf{C}$, we can use the coercivity and the boundedness of the bilinear form $a(\cdot,\cdot)$. In detail, we have, for $v^h = \sum_{i=1}^{n_h} v_i N_i(\mathbf{x}) \in V_0^h \leftrightarrow \underline{v} \in \mathbb{R}^{n_h}$, the following relations:

$$\begin{aligned}\mu_1(\mathbf{B}\underline{v},\underline{v}) &:= \mu_1(v^h,v^h)_V \\ &= \mu_1||v^h||_V^2 \le a(v^h,v^h) = (\mathbf{K}\underline{v},\underline{v}) \le \mu_2||v^h||_V^2 = \mu_2(\mathbf{B}\underline{v},\underline{v}).\end{aligned}$$

Hence, the above defined matrix $\mathbf{B} \in \mathbb{R}^{n_h \times n_h}$ fulfills the spectral equivalence inequalities

$$\mu_1(\mathbf{B}\underline{v},\underline{v}) \le (\mathbf{K}\underline{v},\underline{v}) \le \mu_2(\mathbf{B}\underline{v},\underline{v}) \quad \forall \underline{v} \in \mathbb{R}^{n_h},$$

with positive constants μ_1 and μ_2 which are independent of the discretization parameter h. For our variational problem (3), the inner product $(\cdot,\cdot)_V$ is given by the $H^1(\Omega)$ inner product

$$(u,v)_{H^1(\Omega)} = (u,v)_{L_2(\Omega)} + (\nabla u, \nabla v)_{[L_2(\Omega)]^d}.$$

In the application of the preconditioner, we need the inverse of $\mathbf{B}$ which in this case is usually as expensive as the inversion of the original matrix $\mathbf{K}$. Hence, if an efficient realization of $\mathbf{B}^{-1}$ is not available, we would not have an optimal preconditioned iterative scheme. In boundary element methods the efficient realization of $\mathbf{B}^{-1}$ can be often obtained by operators of inverse orders, see. e.g., Steinbach (2008).

Second idea: Now, let $\underline{u}^{(k)}$ be a current approximation to the solution of the linear system

$$\mathbf{K}\underline{u} = \underline{f}. \tag{32}$$

Hence, $u^{(k)} := \sum_{j=1}^{n_h} u_j^{(k)} N_j \leftrightarrow \underline{u}^{(k)} \in \mathbb{R}^{n_h}$ is an approximation of the solution $u^h \in V_0^h$ of the variational problem

$$a(u^h,v^h) = \ell(v^h) \quad \forall v^h \in V_0^h. \tag{33}$$

In the first attempt, we had to apply the inverse of a matrix which had the same dimension as the original matrix $\mathbf{K}$. Now the idea is to consider only a subspace $W_0^h \subset V_0^h$ with the variational problem

$$\underline{w}^{(k)} \in \mathbb{R}^{n_h} \leftrightarrow w^{(k)} \in W_0 : \; a(w^{(k)},v^h) = \ell(v^h) - a(u^{(k)},v^h) \quad \forall v^h \in W_0^h.$$

Since W_0^h is a subspace of V_0^h, we notice that we have to solve a smaller problem for finding the vector $\underline{w}^{(k)}$. Moreover, we have that

$$a(u^{(k)} + w^{(k)}, v^h) = a(u^{(k)}, v^h) + a(w^{(k)}, v^h)$$
$$= a(u^{(k)}, v^h) + \ell(v^h) - a(u^{(k)}, v^h) = \ell(v^h) \quad \forall v^h \in W_0^h.$$

Hence, for the special case $W_0^h = V_0^h$, we would obtain the exact solution with

$$u = u^{(k)} + w^{(k)} \in V_0^h \qquad \leftrightarrow \qquad \underline{u} = \underline{u}^{(k)} + \underline{w}^{(k)} \in \mathbb{R}^{n_h}.$$

Thus, in this case, $w^{(k)}$ is the correction which has to be added to the current approximation $u^{(k)}$ to obtain the exact solution $u^h \in V_0^h$. If W_0^h is a subspace of V_0^h, we can still interpret the function $w^{(k)} \in W_0^h$ as a correction. This motivates to define the new iterate

$$u^{(k+1)} = u^{(k)} + \tau\, w^{(k)} \in V_0^h \qquad \leftrightarrow \qquad \underline{u}^{(k+1)} = \underline{u}^{(k)} + \tau\, \underline{w}^{(k)} \in \mathbb{R}^{n_h}, \tag{34}$$

where $\tau > 0$ is some positive parameter. We also note that, since W_0^h is only a subspace of V_0^h, we can not correct all components of V_0^h, and the iterative process (34) will in general not converge.

Third idea: In the second attempt, we were not able to correct all components of the error. Now the idea is to split the space V_0^h into several subspaces $W_{0,s}^h \subset V_0^h$ for $s = 1, \ldots, P$, i.e.,

$$V_0^h = \sum_{s=1}^{P} W_{0,s}^h := \left\{ \sum_{s=1}^{P} w_s^h : w_s^h \in W_{0,s}^h \text{ for } s = 1, \ldots, P \right\}.$$

Now, for a given approximation $\underline{u}^{(k)} \in \mathbb{R}^{n_h} \leftrightarrow u^{(k)} \in V_0^h$, we can compute, for each subspace $W_{0,s}^h \subset V_0^h$, $s = 1, \ldots, P$, a subspace correction $w_s^{(k)} \in W_{0,s}^h \leftrightarrow \underline{w}_s^{(k)} \in \mathbb{R}^{n_h}$ as the solution of

$$a(w_s^{(k)}, v_s^h) = \ell(v_s^h) - a(u^{(k)}, v_s^h) \quad \forall v_s^h \in W_{0,s}^h. \tag{35}$$

Now the question is how to combine all the corrections $w_s^{(k)} \in W_{0,s}^h$, $s = 1, \ldots, P$, to obtain a global correction for the current iterate $u^{(k)} \in V_0^h$? There are mainly two possibilities:

- Additive
- Multiplicative

that can also be combined in so-called hybrid versions.

Additive-Schwarz methods For a given approximation $\underline{u}^{(k)} \in \mathbb{R}^{n_h} \leftrightarrow u^{(k)} \in V_0^h$ and the splitting

$$V_0^h = \sum_{s=1}^{P} W_{0,s}^h,$$

we compute the subspace corrections $w_s^{(k)} \in W_{0,s}^h$ for $s = 1, \ldots, P$ as in (35). Then a simple idea to define a global correction is to sum all local corrections together, i.e.,

$$w^{(k)} := \sum_{s=1}^{P} w_s^{(k)} \in V_0^h \quad \leftrightarrow \quad \underline{w}^{(k)} := \sum_{s=1}^{P} \underline{w}_s^{(k)} \in \mathbb{R}^{n_h}.$$

With this global correction we define the next iterate as

$$u^{(k+1)} = u^{(k)} + \tau\, w^{(k)} \in V_0 \quad \leftrightarrow \quad \underline{u}^{(k+1)} = \underline{u}^{(k)} + \tau\, \underline{w}^{(k)} \in \mathbb{R}^{n_h}.$$

Summarizing, we obtain Algorithm 10. Note that all individual corrections in (35) can be computed independently of each other. This is a big advantage if one considers parallel solution algorithms.

Algorithm 10 Additive-Schwarz correction

Given approximation: $\underline{u}^{(k)} \in \mathbb{R}^{n_h} \leftrightarrow u^{(k)} \in V_0^h$
for $s = 1, \ldots, P$ **do**
  Find $w_s^{(k)} \in W_{0,s}^h : a(w_s^{(k)}, v_s^h) = \ell(v_s^h) - a(u^{(k)}, v_s^h)\ \forall v_s^h \in W_{0,s}^h$
end for
Compute correction: $w^{(k)} = \sum_{s=1}^{P} w_s^{(k)}$
Compute update: $u^{(k+1)} = u^{(k)} + \tau\, w^{(k)} \in V_0^h \leftrightarrow \underline{u}^{(k+1)} \in \mathbb{R}^{n_h}$

Example 5.2 (*Jacobi method*) For the discrete function space

$$V_0^h = \operatorname{span}\{N_s\}_{s=1}^{n_h},$$

we consider the one-dimensional subspaces

$$W_{0,s}^h := \operatorname{span}\{N_s\} \quad \text{for } s = 1, \ldots, n_h.$$

Hence each subspace $W_{0,s}^h$ is spanned by only one basis function $N_s \in V_0^h$. Thus, we clearly have

$$V_0^h = \sum_{s=1}^{P} W_{0,s}^h \quad \text{with } P = n_h.$$

Then the additive correction is given by

$$w^{(k)} = \sum_{s=1}^{n_h} w_s^{(k)} = \sum_{s=1}^{n_h} w_s N_s \quad \leftrightarrow \quad \underline{w}^{(k)} = [w_s]_{s=1}^{n_h} \in \mathbb{R}^{n_h},$$

where the corrections $w_s^{(k)} \in W_{0,s}^h = \text{span}\{N_s\} \leftrightarrow w_s \in \mathbb{R}$ are given by the variational problem

$$a(w_s^{(k)}, v_s^h) = \ell(v_s^h) - a(u^{(k)}, v_s^h) \quad \forall v_s^h \in W_{0,s}.$$

This is equivalent in finding $w_s \in \mathbb{R}$ such that

$$a(N_s, N_s) w_s = \ell(N_s) - a(u^{(k)}, N_s).$$

By using the definition of the matrix $\mathbf{K} = \left[a(N_j, N_i)\right]_{i,j=1,\dots,n_h}$ and the vector $\underline{f} = [\ell(N_i)]_{i=1,\dots,n_h}$, we find that

$$K_{ss} w_s = f_s - \left[\mathbf{K}\underline{u}^{(k)}\right]_s = \left[\underline{f} - \mathbf{K}\underline{u}^{(k)}\right]_s.$$

Summarizing, we have

$$\underline{w}^{(k)} = [w_s]_{s=1}^{n_h} \in \mathbb{R}^{n_h} \quad \text{with} \quad w_s = K_{ss}^{-1} \left[\underline{f} - \mathbf{K}\underline{u}^{(k)}\right]_s.$$

Hence, the correction is given by

$$\underline{w}^{(k)} = \mathbf{D}^{-1} \left[\underline{f} - \mathbf{K}\underline{u}^{(k)}\right] \quad \text{with} \quad \mathbf{D} := \text{diag}(\mathbf{K}).$$

Then the next iterate is obtained by

$$\underline{u}^{(k+1)} = \underline{u}^{(k)} + \tau \underline{w}^{(k)} = \underline{u}^{(k)} + \tau \mathbf{D}^{-1} \left[\underline{f} - \mathbf{K}\underline{u}^{(k)}\right]. \tag{36}$$

This scheme (36) is nothing else then the Richardson method with the "preconditioner" $\mathbf{D}^{-1}$ or the so called damped Jacobi method, compare also Algorithm 5. Note that this scheme is not robust with respect to the discretization parameter h, see Example 2.1.

Algorithm 11 Multiplicative-Schwarz correction

Given approximation: $\underline{u}^{(k)} \in \mathbb{R}^{n_h} \leftrightarrow u^{(k)} \in V_0^h$
Define: $u_0^{(k)} := u^{(k)}$
for $s = 1, \dots, P$ **do**
  Find $w_s^{(k)} \in W_{0,s}^h : a(w_s^{(k)}, v_s^h) = \ell(v_s^h) - a(u_{s-1}^{(k)}, v_s^h)\ \forall v_s^h \in W_{0,s}^h$
  $u_s^{(k)} = u_{s-1}^{(k)} + w_s^{(k)}$
end for
Final update: $u^{(k+1)} = u_P^{(k)} \in V_0^h \leftrightarrow \underline{u}^{(k+1)} \in \mathbb{R}^{n_h}$

Multiplicative-Schwarz methods For a given approximation $\underline{u}^{(k)} \in \mathbb{R}^{n_h} \leftrightarrow u^{(k)} \in V_0^h$, we again consider a splitting

$$V_0^h = \sum_{s=1}^{P} W_{0,s}^h.$$

Now the idea is to compute one correction after the other, and apply the correction immediately for each subspace. This results in the Algorithm 11.

For the multiplicative scheme, we sum up the following remarks:

- The ordering of the subspaces $W_{0,s}^h \subset V_0^h$ plays a role.
- The application of each correction is a sequential process.
- A damping parameter for each correction can be introduced.
- One can also combine multiplicative and additive correction steps to obtain parallel methods → for example like Multigrid methods (multiplicative over the levels and additive or multiplicative in each level), see Sect. 6.

Example 5.3 (*Gauss–Seidel method*) As in Example 5.2, we consider the splitting of V_0^h into the one-dimensional subspaces

$$V_0^h = \sum_{s=1}^{P} W_{0,s}^h, \quad \text{with } W_{0,s}^h := \text{span}\{N_s\} \quad \text{for } s = 1, \ldots, n_h = P.$$

Then, according to Algorithm 11, the global multiplicative correction $w^{(k)} \in V_0^h \leftrightarrow \underline{w}^{(k)} \in \mathbb{R}^{n_h}$ is given by

$$\underline{w}^{(k)} = \mathbf{L}^{-1}\left[\underline{f} - \mathbf{K}\underline{u}^{(k)}\right],$$

where $\mathbf{L}$ is the lower triangular matrix of $\mathbf{K}$. Hence the next iterate is given by

$$\underline{u}^{(k+1)} = \underline{u}^{(k)} + \mathbf{L}^{-1}\left[\underline{f} - \mathbf{K}\underline{u}^{(k)}\right]. \tag{37}$$

The obtained scheme (37) is the Gauss–Seidel iteration as already explained in Algorithm 6. Note that this method is also not robust with respect to the discretization parameter h, see Sect. 4.

Multilevel diagonal scaling Based on the subspace corrections from above, we will now derive an additive scheme which will result in an efficient preconditioner for our model problem. The idea is to apply the subspace corrections from Example 5.2 to a sequence of nested spaces

$$V_0^0 \subset V_0^1 \subset \cdots \subset V_0^L = V_0^h, \quad \text{with } \dim(V_0^\ell) = n_\ell \text{ for } \ell = 0, \ldots, L.$$

Note, to simplify the notation, we will skip from now on the index h in the definition of the discrete spaces and functions. On the finest space $V_0^L = V_0^h$, we want to solve our linear system

$$\mathbf{K}\underline{u} = \underline{f}.$$

We can obtain such a nested sequence of spaces by constructing a nested sequence of finite element meshes (for example by applying uniform refinement to the finite element mesh several times). A 1d illustration of such a nested sequence of finite element meshes is given in Fig. 2. For each space V_0^ℓ, $\ell = 0, \dots, L$, we introduce the basis functions

$$V_0^\ell = \operatorname{span}\{N_j^\ell\}_{j=1}^{n_\ell} \quad \text{for } \ell = 0, \dots, L.$$

Using the subspace corrections from Example 5.2 for each level $\ell = 0, \dots, L$, we arrive at the splitting

$$V_0^\ell = \sum_{i=1}^{n_\ell} W_{0,i}^\ell, \quad \text{with } W_{0,i}^\ell := \operatorname{span}\{N_i^\ell\} \quad \text{for } i = 1, \dots, n_\ell$$

for each space V_0^ℓ. This results in the overall subspace decomposition

$$V_0^L = \sum_{\ell=0}^{L} V_0^\ell = \sum_{\ell=0}^{L} \sum_{i=1}^{n_\ell} W_{0,i}^\ell.$$

Now, for a given approximation $u^{(k)} \in V_0^L$, the additive correction is then given by

$$w^{(k)} = \sum_{\ell=1}^{L} \sum_{i=1}^{n_\ell} w_i^\ell N_i^\ell = \sum_{\ell=1}^{L} w^\ell, \quad \text{with } w^\ell := \sum_{i=1}^{n_\ell} w_i^\ell N_i^\ell,$$

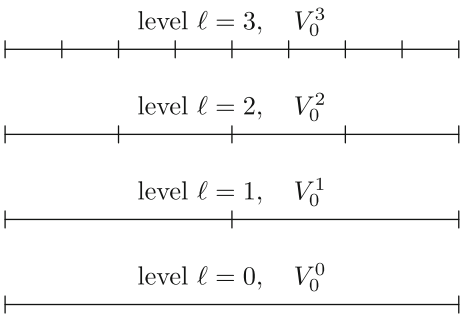


Fig. 2 A nested sequence of 1d finite element meshes

where the coefficients $w_i^\ell \in \mathbb{R}$ are given by the subspace correction equations

$$a(N_i^\ell, N_i^\ell) w_i^\ell = \ell(N_i^\ell) - a(u^{(k)}, N_i^\ell) =: \langle R^\ell, N_i^\ell \rangle =: \left[\underline{r}^\ell\right]_i .$$

Hence, we have

$$\underline{w}^\ell = \mathbf{D}_\ell^{-1} \underline{r}^\ell \quad \text{with } \mathbf{D}_\ell := \operatorname{diag}(\mathbf{K}_\ell), \text{ where } \mathbf{K}_\ell := [a(N_j^\ell, N_i^\ell)]_{i,j=1}^{n_\ell}.$$

With the computed correction $w^{(k)}$ we then can compute the next iterate as

$$u^{(k+1)} = u^{(k)} + \tau\, w^{(k)}.$$

Summarizing the above computations results in Algorithm 12.

Algorithm 12 Multilevel diagonal scaling (MDS)

Given approximation: $\underline{u}^{(k)} \in \mathbb{R}^{n_h} \leftrightarrow u^{(k)} \in V_0^h$
For each level $\ell = 0, \ldots, L$ compute the residual

$$\underline{r}^\ell := \left[\ell(N_i^\ell) - a(u^{(k)}, N_i^\ell)\right]_{i=1}^{n_\ell}$$

Apply diagonal scaling for each level $\ell = 0, \ldots, L$

$$\underline{w}^\ell = \mathbf{D}_\ell^{-1} \underline{r}^\ell \quad \leftrightarrow \quad w^\ell \in V_0^\ell$$

Sum up all corrections

$$w^{(k)} = \sum_{\ell=0}^{L} w^\ell \in V_0 \quad \leftrightarrow \quad \underline{w}^{(k)} \in \mathbb{R}^{n_h}$$

Compute update: $\underline{u}^{(k+1)} = \underline{u}^{(k)} + \alpha\, \underline{w}^{(k)}$

We notice, that every computation in Algorithm 12 is linear with respect to the residual

$$\underline{r}^{(k)} := \underline{f} - \mathbf{K}\underline{u}^{(k)}.$$

Hence, there exists a matrix

$$\mathbf{C}_{\text{MDS}}{}^{-1} : \mathbb{R}^{n_h} \to \mathbb{R}^{n_h},$$

such that

$$\underline{w}^{(k)} = \mathbf{C}_{\text{MDS}}{}^{-1} \underline{r}^{(k)} = \mathbf{C}_{\text{MDS}}{}^{-1} \left[\underline{f} - \mathbf{K}\underline{u}^{(k)}\right].$$

Thus, the scheme given in Algorithm 12 results in the preconditioned Richardson method

$$\underline{u}^{(k+1)} = \underline{u}^{(k)} + \tau\, \mathbf{C}_{\mathrm{MDS}}{}^{-1} \left[\underline{f} - \mathbf{K}\underline{u}^{(k)} \right] \quad \text{for } k = 0, 1, \ldots$$

with preconditioner $\mathbf{C}_{\mathrm{MDS}}{}^{-1}$. Of course, for an efficient implementation, the matrix $\mathbf{C}_{\mathrm{MDS}}{}^{-1}$ will not be computed. Only the action onto the residual, which is given in Algorithm 12, will be implemented. We also notice that one iteration of the multilevel diagonal scaling is of optimal complexity $\mathcal{O}(n_h)$ and the usual transfer operators are used between different levels, see also Sect. 6. Moreover, the preconditioner $\mathbf{C}_{\mathrm{MDS}}{}^{-1}$ can be used in other iterative methods like the preconditioned Conjugate Gradient method presented in Algorithm 9. In the next theorem, we present bounds for the condition number of the preconditioned matrix $\mathbf{C}_{\mathrm{MDS}}{}^{-1}\mathbf{K}$.

Theorem 5.4 *For the multilevel diagonal scaling preconditioner* $\mathbf{C}_{\mathrm{MDS}}$*, one can show the spectral equivalence estimates*

$$\mu_1(\mathbf{C}_{\mathrm{MDS}}\underline{v}, \underline{v}) \leq (\mathbf{K}\underline{v}, \underline{v}) \leq \mu_2(\mathbf{C}_{\mathrm{MDS}}\underline{v}, \underline{v}) \quad \forall \underline{v} \in \mathbb{R}^{n_h},$$

with constants μ_1 *and* μ_2 *that are independent of* h *(only* $\log(h^{-1})$*).*

The multilevel diagonal scaling preconditioner $\mathbf{C}_{\mathrm{MDS}}$ was introduced by Zhang (1992) where one can also find the proof of Theorem 5.4, see also Oswald (1999) for the case of jumping coefficients where logs can appear. The MDS preconditioner generalizes the BPX preconditioner that was earlier introduced by Bramble et al. (1990).

6 Multigrid Methods

6.1 Motivation

We will motivate the multigrid methods by first looking at a very simple example, namely the 1d Poisson problem. We consider the computational domain $\Omega = (0, 1)$ uniformly decomposed into elements with mesh size h. For the discrete function space V_0^h, we consider continuous and piecewise linear functions. So we arrive at the discrete variational problem

$$\text{Find } u^h \in V_0^h : \quad a(u^h, v^h) = \ell(v^h) \quad \forall v^h \in V_0^h,$$

with

$$a(u^h, v^h) = \int_0^1 \frac{\mathrm{d}u^h}{\mathrm{d}x}(x)\frac{\mathrm{d}v^h}{\mathrm{d}x}(x)\mathrm{d}x \quad \text{and} \quad \ell(v^h) = \int_0^1 f(x)v^h(x)\mathrm{d}x.$$

Therefore, we have to solve the linear system

$$\mathbf{K}\underline{u} = \underline{f}, \tag{38}$$

with

$$\mathbf{K} = \frac{1}{h}\begin{pmatrix} 2 & -1 & & \\ -1 & 2 & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{pmatrix} \quad \text{and} \quad \underline{f} = \left[\int_0^1 f(x)N_i(x)\mathrm{d}x\right]_{i=1}^{n_h}.$$

Now we apply the damped Jacobi method, see Algorithm 5 or Eq. (36), to the linear system (38) for $h = 2^{-7}$ with the special right hand side $\underline{f} = \underline{0}$ and a random initial guess $\underline{u}^{(0)} = [\mathrm{rand}(0,1)]_{j=1}^{n_h}$ with values between zero and one. Hence the exact solution of (38) is given by $\underline{u} = \underline{0}$ and each iterate $\underline{u}^{(k)}$ of the damped Jacobi scheme is equal to the error $\underline{e}^{(k)} = \underline{u}^{(k)} - \underline{u}$. In Fig. 3, we plotted the error for different damping parameters $\tau \in \{1, \frac{2}{3}\}$ and different iterations.

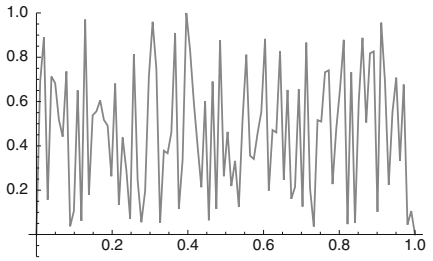
For the optimal parameter $\tau_{opt} = 1$, see Example 4.6, we observe that the error is reduced very slowly, and that the high oscillations from the initial error still occur after 30 iterations, whereas, for a parameter $\tau = \frac{2}{3}$, which is not optimal in the sence of the fastest convergence, we observe that the high oscillations of the error are reduced very fast. We say that the error is getting smoother. To explain this behaviour we look at the Fourier expansion of the initial error $\underline{e}^{(0)}$. We recall the eigenvectors and eigenvalues for the matrix $\mathbf{K} \in \mathbb{R}^{(n-1)\times(n-1)}$, see Example 2.1, i.e.

$$\mathbf{K}\,\underline{\varphi}_i = \lambda_i \underline{\varphi}_i \quad \text{with } \lambda_i = \frac{4}{h}\sin^2\left(\frac{i\,\pi}{2n}\right) \text{ and } \underline{\varphi}_i = \left[\sqrt{2n}\,\sin(ik\pi h)\right]_{k=1}^{n-1}.$$

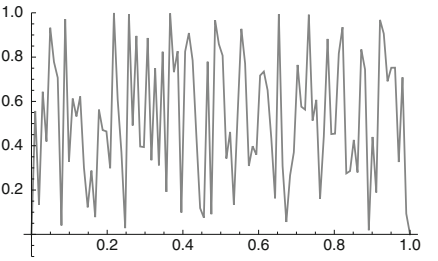
Hence, we can write the initial error as

$$\underline{e}^{(0)} := \underline{u}^{(0)} - \underline{u} = \sum_{i=1}^{n-1} \alpha_i \underline{\varphi}_i,$$

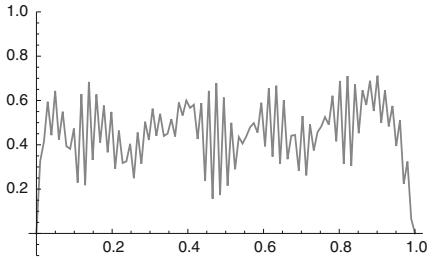
with coefficients $\alpha_i = (\underline{e}^{(0)}, \underline{\varphi}_i)$ for $i = 1, \ldots, n-1$. We then obtain, for the error iteration scheme (see also (24)), the identity



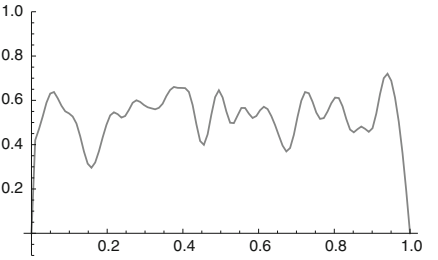
Initial error for $\tau = 1$.



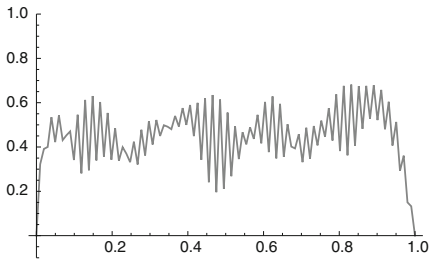
Initial error for $\tau = \frac{2}{3}$.



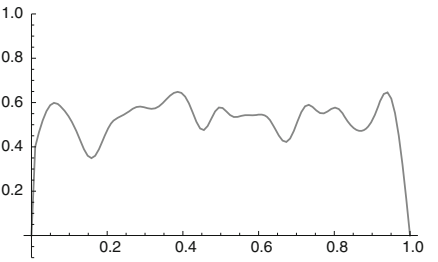
Iteration $k = 5$ for $\tau = 1$.



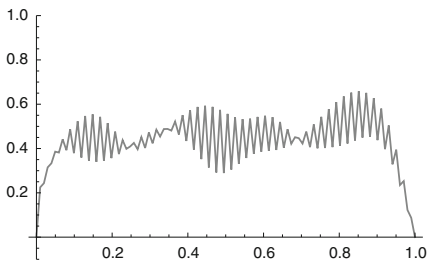
Iteration $k = 5$ for $\tau = \frac{2}{3}$.



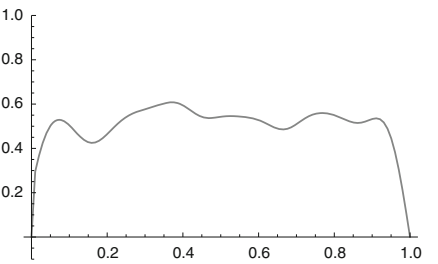
Iteration $k = 10$ for $\tau = 1$.



Iteration $k = 10$ for $\tau = \frac{2}{3}$.



Iteration $k = 30$ for $\tau = 1$.



Iteration $k = 30$ for $\tau = \frac{2}{3}$.

Fig. 3 Errors of the damped Jacobi scheme for different iterations and two different damping parameters $\tau \in \{1, \frac{2}{3}\}$

$$\underline{e}^{(k+1)} = \mathbf{E}\,\underline{e}^{(k)} = \left[I - \tau \mathbf{D}^{-1}\,\mathbf{K}\right]\underline{e}^{(k)} = \left[I - \tau \mathbf{D}^{-1}\,\mathbf{K}\right]^k \sum_{i=1}^{n_h} \alpha_i \underline{\varphi}_i$$
$$= \sum_{i=1}^{n_h} \alpha_i \left[1 - \tau \frac{h}{2}\lambda_i\right]^k \underline{\varphi}_i = \sum_{i=1}^{n_h} \alpha_i \left[1 - 2\tau \sin^2\left(\frac{i\,\pi}{2n}\right)\right]^k \underline{\varphi}_i.$$

We now estimate the expression

$$\left|1 - 2\tau \sin^2\left(\frac{i\,\pi}{2n}\right)\right|$$

for different $i = 1, \ldots, n-1$. For even n, we start with $i = 1, \ldots, \frac{n}{2}$, i.e., the low oscillating error components, and obtain

$$\left|1 - 2\tau \sin^2\left(\frac{i\,\pi}{2n}\right)\right| \le \max\left\{\left|1 - 2\tau \sin^2\left(\frac{\pi}{2n}\right)\right|, |1-\tau|\right\}$$
$$\approx \left|1 - \tau\frac{\pi^2}{2}h^2\right| \approx 1,$$

for small $h \ll 1$. For the high oscillating error components $i = \frac{n}{2}, \ldots, n-1$, we obtain

$$\left|1 - 2\tau \sin^2\left(\frac{i\,\pi}{2n}\right)\right| \le \max\ \{|1-\tau|, |1-2\tau|\} = \frac{1}{3} \quad \text{for } \tau^*_{opt} = \frac{2}{3}.$$

Hence with the choice $\tau^*_{opt} = \frac{2}{3}$ the high oscillating part of the error is reduced in each iteration by a factor of $\frac{1}{3}$, whereas the low oscillating part of the error is almost not reduced for very small h. This exactly explains the behaviour observed in Fig. 3. We also mention that other iterative schemes like the Gauss–Seidel iteration have a similar behaviour.

6.2 Two-Grid Cycle

We observed that simple iterative methods like the damped Jacobi method or the Gauss–Seidel iteration reduce the high oscillatory part of the error very fast. Hence,

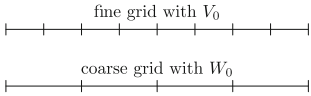


Fig. 4 Fine and coarse grids for a 1d-problem

the error function $e^{(k)} = u^{(k)} - u$ is getting smoother and smoother. Therefore, the best possible correction is also a smooth function, which in general can be well approximated by a coarser grid. This observation now motivates to apply a subspace correction after a smoothing procedure, where the subspace $W_0^h \subset V_0^h$ is comming from a coarser grid, see also Fig. 4 for a 1d problem. For a smoothed approximation $u^{(k)} \in V_0^h$, we can compute the subspace correction $w^{(k)} \in W_0^h \leftrightarrow \underline{w}^{(k)} \in \mathbb{R}^{n_h}$ by using the subspace correction equation as introduced in Sect. 5, i.e.,

$$a(w^{(k)}, v^h) = \ell(v^h) - a(u^{(k)}, v^h) \quad \forall v^h \in W_0^h. \tag{39}$$

Considering again a basis $\{N_i^C\}_{i=1}^{n_C}$ for the subspace W_0^h, $\dim(W_0^h) = n_C$, the discrete problem (39) is equivalent to a linear system

$$\mathbf{K}_C \, \underline{w}_C^{(k)} = \underline{r}_C^{(k)}, \tag{40}$$

with

$$\mathbf{K}_C := \left[a(N_j^C, N_i^C)\right]_{i,j=1}^{n_C} \quad \text{and} \quad \underline{r}_C^{(k)} := \left[\ell(N_i^C) - a(u^{(k)}, N_i^C)\right]_{i=1}^{n_C}.$$

So we have to find out the connection between the coarse grid coefficient vector $\underline{w}_C^{(k)} \in \mathbb{R}^{n_C}$ and the fine grid coefficient vector $\underline{w}^{(k)} \in \mathbb{R}^{n_h}$? For any $w^{(k)} \in W_0^h \subset V_0^h$, we have the two possible representations

$$w^{(k)} = \sum_{i=1}^{n_C} w_i^C N_i^C \quad \text{or} \quad w^{(k)} = \sum_{j=1}^{n_h} w_j N_j.$$

Therefore, any coarse grid basis function $N_i^C \in W_0^h \subset V_0^h$ for $i = 1, \ldots, n_C$ can also be written by using the fine grid basis functions

$$N_i^C = \sum_{j=1}^{n_h} P[j,i] N_j, \tag{41}$$

with some coefficients $P[j,i] \in \mathbb{R}$ for $j = 1, \ldots, n_h$. Thus, we further obtain

$$\begin{aligned} w^{(k)} &= \sum_{i=1}^{n_C} w_i^C N_i^C = \sum_{i=1}^{n_C} w_i^C \left[\sum_{j=1}^{n_h} P[j,i] N_j \right] \\ &= \sum_{j=1}^{n_h} \left[\sum_{i=1}^{n_C} P[j,i] w_i^C \right] N_j = \sum_{j=1}^{n_h} \left[\mathbf{P} \underline{w}_C^{(k)} \right]_j N_j. \end{aligned}$$

Since the coefficient with respect to a basis are uniquely defined, we obtain the important relation

$$\underline{w}^{(k)} = \mathbf{P}\underline{w}_C^{(k)}, \tag{42}$$

where $\mathbf{P} \in \mathbb{R}^{n_h \times n_C}$ is the so called prolongation matrix. For solving the coarse grid problem (40), we have to compute the coarse grid residual

$$\underline{r}_C^{(k)} \in \mathbb{R}^{n_C} \quad \leftrightarrow \quad \langle R^{(k)}, v^h \rangle := \ell(v^h) - a(u^{(k)}, v^h) \quad \forall v^h \in W_0^h.$$

Using again the representation (41), we can write the coarse grid residual as

$$\begin{aligned} \underline{r}_C^{(k)}[i] &= \langle R^{(k)}, N_i^C \rangle = \langle R^{(k)}, \sum_{j=1}^{n_h} P[j,i] N_j \rangle \\ &= \sum_{j=1}^{n_h} P[j,i] \langle R^{(k)}, N_j \rangle = \sum_{j=1}^{n_h} P[j,i] \underline{r}^{(k)}[j] = \left[\mathbf{P}^\top \underline{r}^{(k)}\right]_i, \end{aligned}$$

where $\underline{r}^{(k)} = \underline{f} - \mathbf{K}\underline{u}^{(k)}$ is the fine grid residual. Hence, we have shown that

$$\underline{r}_C^{(k)} = \mathbf{P}^\top \underline{r}^{(k)} =: \mathbf{R}\,\underline{r}^{(k)}, \tag{43}$$

where $\mathbf{R} := \mathbf{P}^\top \in \mathbb{R}^{n_C \times n_h}$ is the so called restriction matrix. Note that the prolongation matrix $\mathbf{P}$ corresponds to the basis transformation with respect to the trial space and the restriction matrix is the transposed prolongation matrix which corresponds to the basis transformation coming from the test space. Since in our case the trial and test space are the same, we have $\mathbf{R} = \mathbf{P}^\top$.

Example 6.1 To illustrate how the prolongation and restriction matrices look like, we study the simple 1d-problem from Sect. 6.1. In Fig. 5, fine and coarse grid basis functions are plotted. If we now want to represent the coarse grid basis function N_2^C, we obtain

$$N_2^C = \frac{1}{2}N_3 + 1N_4 + \frac{1}{2}N_5$$

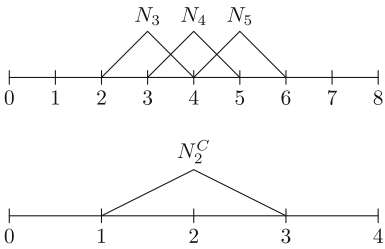


Fig. 5 Fine and coarse grid basis functions for a 1d-problem

by interpolation. Hence, we have

$$P[3,2] = \frac{1}{2}, \qquad P[4,2] = 1, \qquad P[5,2] = \frac{1}{2}.$$

Thus, for the example illustrated in Fig. 5, we have the following prolongation and restriction matrices

$$\mathbf{P} = \begin{pmatrix} \frac{1}{2} & & \\ 1 & & \\ \frac{1}{2} & \frac{1}{2} & \\ & 1 & \\ & \frac{1}{2} & \frac{1}{2} \\ & & 1 \\ & & \frac{1}{2} \end{pmatrix} \qquad \text{and} \qquad \mathbf{R} = \mathbf{P}^\top.$$

Hence, the prolongation and restriction matrices can be represented by sparse matrices, which means that the grid transfer operations are of optimal complexity.

Algorithm 13 Two-grid cycle

Given initial approximation $\underline{u}^{(k)}$ and right hand side $\underline{f}$
Apply pre-smoothing: $\underline{u}^{(k)} = S^{\nu_1}(\underline{u}^{(k)}, \underline{f})$
Compute defect: $\underline{d}^{(k)} = \underline{f} - \mathbf{K}\,\underline{u}^{(k)}$
Restriction: $\underline{d}_C^{(k)} = \mathbf{R}\,\underline{d}^{(k)}$
Solve coarse grid problem: $\mathbf{K}_C \underline{w}_C^{(k)} = \underline{d}_C^{(k)}$
Prolongation: $\underline{w}^{(k)} = \mathbf{P}\,\underline{w}_C^{(k)}$
Correction: $\underline{u}^{(k)} = \underline{u}^{(k)} + \underline{w}^{(k)}$
Apply post-smoothing: $\underline{u}^{(k)} = S^{\nu_2}(\underline{u}^{(k)}, \underline{f})$

Summarizing, the above computations gives the so called two-grid cycle, see Algorithm 13. Note that $\underline{u}^{(k)} = S^{\nu}(\underline{u}^{(k)}, \underline{f})$ means the application of ν-steps of a smoothing procedure like the damped Jacobi method or the Gauss–Seidel scheme for example. Now the question arises under which conditions the iterative application of two-grid cycle is convergent? In literature there are two possible analysis tools available

- additive splitting and Fourier analysis using eigenvalues and eigenvectors of $\mathbf{K}$, and
- multiplicative splitting.

For both analysis tools, we need the error iteration matrix for the two-grid cycle. Since we already know the error iteration matrix $\mathbf{E} = \left[I - \tau \mathbf{C}^{-1}\,\mathbf{K}\right]$ for the smoother,

see (24), we only have to investigate the error iteration matrix for the coarse grid correction step. For this, let $\underline{u}^{(k)}$ be the approximation obtained after the smoothing procedure with error $\underline{e}^{(k)} = \underline{u}^{(k)} - \underline{u}$, and let $\underline{w}^{(k)}$ be the coarse grid correction. Then the error after the coarse grid correction step is given by

$$\begin{aligned}
\underline{e}_{\text{cor}}^{(k)} &:= \left(\underline{u}^{(k)} + \underline{w}^{(k)}\right) - \underline{u} = \underline{e}^{(k)} + \underline{w}^{(k)} = \underline{e}^{(k)} + \mathbf{P}\,\underline{w}_C^{(k)} \\
&= \underline{e}^{(k)} + \mathbf{P}\,\mathbf{K}_C^{-1}\,\underline{d}_C^{(k)} \\
&= \underline{e}^{(k)} + \mathbf{P}\,\mathbf{K}_C^{-1}\,\mathbf{R}\,\underline{d}^{(k)} = \underline{e}^{(k)} + \mathbf{P}\,\mathbf{K}_C^{-1}\,\mathbf{R}\left[\underline{f} - \mathbf{K}\,\underline{u}^{(k)}\right] \\
&= \underline{e}^{(k)} - \mathbf{P}\,\mathbf{K}_C^{-1}\,\mathbf{R}\,\mathbf{K}\,\underline{e}^{(k)} = \left[I - \mathbf{P}\,\mathbf{K}_C^{-1}\,\mathbf{R}\,\mathbf{K}\right]\underline{e}^{(k)} \\
&=: \mathbf{T}\underline{e}^{(k)}.
\end{aligned}$$

Thus, the error iteration matrix for the coarse grid correction step is given by

$$\mathbf{T} = I - \mathbf{P}\,\mathbf{K}_C^{-1}\,\mathbf{R}\,\mathbf{K}.$$

Hence, the error of the two-grid cycle can be represented in the form

$$\underline{e}_{\text{tg}}^{(k+1)} = \mathbf{E}^{\nu_2}\,\mathbf{T}\,\mathbf{E}^{\nu_1}\,\underline{e}_{\text{tg}}^{(k)} = \mathbf{M}\,\underline{e}_{\text{tg}}^{(k)} = \mathbf{M}^k\,\underline{e}_{\text{tg}}^{(0)},$$

with the error iteration matrix $\mathbf{M} := \mathbf{E}^{\nu_2}\,\mathbf{T}\,\mathbf{E}^{\nu_1} = \mathbf{E}^{\nu_2}\left[I - \mathbf{P}\,\mathbf{K}_C^{-1}\,\mathbf{R}\,\mathbf{K}\right]\mathbf{E}^{\nu_1}$.

First attempt A first simple idea to estimate the error iteration matrix $\mathbf{M} = \mathbf{E}^{\nu_2}\,\mathbf{T}\,\mathbf{E}^{\nu_1}$ is given by the estimate

$$||\mathbf{M}|| = ||\mathbf{E}^{\nu_2}\,\mathbf{T}\,\mathbf{E}^{\nu_1}|| \le ||\mathbf{T}||\;||\mathbf{E}||^{\nu_1+\nu_2}. \tag{44}$$

For our model problem, we know that

$$||\mathbf{E}||^{\nu} = \left[1 - \mathcal{O}(h^2)\right]^{\nu} \to 0 \quad \text{for } \nu \to \infty.$$

converges very slowly to 0. On the other side, we have

$$\begin{aligned}
||\mathbf{T}|| &= \sup_{0\neq \underline{v}\in\mathbb{R}^{n_h}} \frac{||\mathbf{T}\,\underline{v}||}{||\underline{v}||} = \sup_{0\neq \underline{v}\in\mathbb{R}^{n_h}} \frac{||\left[I - \mathbf{P}\,\mathbf{K}_C^{-1}\,\mathbf{R}\,\mathbf{K}\right]\underline{v}||}{||\underline{v}||} \\
&\ge \sup_{\substack{0\neq \underline{v}\in\mathbb{R}^{n_h} \\ \mathbf{K}\,\underline{v}\in\ker(\mathbf{R})}} \frac{||\left[I - \mathbf{P}\,\mathbf{K}_C^{-1}\,\mathbf{R}\,\mathbf{K}\right]\underline{v}||}{||\underline{v}||} = 1.
\end{aligned}$$

Thus, the simple splitting (44) is not successful.

Second attempt We here use more the structure of $\mathbf{T}$ and the multiplicative splitting

$$||\mathbf{T}\,\mathbf{E}^{\nu}|| = ||\mathbf{T}\,\mathbf{K}^{-1}\,\mathbf{K}\,\mathbf{E}^{\nu}|| \le ||\mathbf{T}\,\mathbf{K}^{-1}||\;||\mathbf{K}\,\mathbf{E}^{\nu}||.$$

If we can now show the following two properties

- Approximation property

$$||\mathbf{T}\,\mathbf{K}^{-1}|| \leq c\,h^{\delta}, \quad \text{with some } \delta > 0$$

- Smoothing property

$$||\mathbf{K}\,\mathbf{E}^{\nu}|| \leq \eta(\nu) h^{-\delta} \quad \text{with } \eta(\nu) \to 0 \quad \text{as } \nu \to \infty,$$

then we have convergence for the two-grid cycle. Indeed, we have

$$||\mathbf{M}|| \leq ||\mathbf{T}\,\mathbf{E}^{\nu}|| \leq ||\mathbf{T}\,\mathbf{K}^{-1}||\,||\mathbf{K}\,\mathbf{E}^{\nu}|| \leq c\,\eta(\nu) < 1$$

for $\nu \in \mathbb{N}$ large enough. Under some regularity assumptions, one can show the approximation and smoothing properties for the d-dimensional Poisson problem and for more general elliptic boundary value problems, see Hackbusch (1985).

6.3 Multigrid Cycle

What should we do if the coarse problem is still too large to solve it by means of a direct method efficiently? The idea is to approximate the exact solution of the coarse grid problem by another or several two-grid cycles and repeat this procedure recursively. So we again need a hierarchy of grids as it was the case for the multilevel diagonal scaling preconditioner, see Algorithm 12. For each level $\ell = 0, 1, \ldots, L$, we use the following notations

- System matrix $\mathbf{K}_\ell$, solution vector $\underline{u}_\ell$ and right hand side vector $\underline{f}_\ell$,
- Restriction matrix $\mathbf{R}_\ell$ acting between level ℓ and level $\ell - 1$,
- Prolongation matrix $\mathbf{P}_\ell$ acting between level $\ell - 1$ and level ℓ.

Using these notations, we immediately obtain the multigrid cycle given recursively in Algorithm 14, where we skipped the iteration index for a simpler notation. Moreover, γ denotes the cycle-index. A usual choice is $\gamma = 1$ (V-cycle) or $\gamma = 2$ (W-cycle). Of course a V-cycle is the cheapest cycle but the analysis is more difficult in general. For a W-cycle the analysis is easier but it leads to a more expensive method. The multigrid method was introduced by Fedorenko (1961) who also provided the first W-cycle analysis based on the additive splitting in Fedorenko (1964), see also Bakhvalov (1966) for a more general setting. The first V-cycle proof goes back to Braess and Hackbusch (1983).

Algorithm 14 MGCycle

Require: $\underline{u}_\ell$, $\underline{f}_\ell$
  if $\ell = 0$ **then**
    Coarse grid solver: $\underline{u}_\ell = \mathbf{K}_\ell^{-1}\,\underline{f}_\ell$
  else
    Pre-smoothing: $\underline{u}_\ell = S_\ell^{\nu_1}(\underline{u}_\ell, \underline{f}_\ell)$
    Compute defect: $\underline{d}_\ell = \underline{f}_\ell - \mathbf{K}_\ell\,\underline{u}_\ell$
    Restriction: $\underline{d}_{\ell-1} = \mathbf{R}_\ell\,\underline{d}_\ell$
    Initialize: $\underline{w}_{\ell-1} = 0$
    for $i = 1, \ldots, \gamma$ **do**
      MGCycle($\underline{w}_{\ell-1}$,$\underline{d}_{\ell-1}$)
    end for
    Prolongation: $\underline{w}_\ell = \mathbf{P}_\ell\,\underline{w}_{\ell-1}$
    Correction: $\underline{u}_\ell = \underline{u}_\ell + \underline{w}_\ell$
    Post-smoothing: $\underline{u}_\ell = S_\ell^{\nu_2}(\underline{u}_\ell, \underline{f}_\ell)$
  end if

Algorithm 15 Full multigrid cycle

Coarse problem: $\underline{u}_0 = \mathbf{K}_0^{-1}\,\underline{f}_0$
for $\ell = 1, \ldots, L$ **do**
  Prolongate: $\underline{u}_\ell = \mathbf{P}_\ell\,\underline{u}_{\ell-1}$
  Apply MG-cycle MGCycle($\underline{u}_\ell$, $\underline{f}_\ell$) until discretization error is reached
end for

Bakhvalov (1966) also proposed to use multigrid cycles in the framework of a nested iteration procedure starting from the coarsest level to the finest level. This procedure is now called full multigrid cycle that is presented in Algorithm 15. It is possible to arrange the full multigrid cycle in such a way that it produces an approximation $\widetilde{\underline{u}}_\ell$ that differs from the exact solution in the order of the discretization error. Since one can obtain this with a constant number of nested iterations at all levels, the arithmetical complexity of the full multigrid method is proportional to the number of unknowns on the finest grid, i.e., this method is asymptotically optimal. The full multigrid cycle has advantages when considering adaptive mesh refinement or when non-linear problems have to be solved.

The reader who is interested in the numerical analysis of multigrid methods is referred to the monographs by Hackbusch (1985) and Bramble (1993).

6.4 Time-Parallel Multigrid

The multigrid method can also be used to solve time dependent problems simultaneously. For this we consider the simple scalar ordinary differential equation

$$\begin{aligned} \frac{\mathrm{d}u}{\mathrm{d}t}(t) + \lambda\, u(t) &= f(t), \qquad \text{for } t \in (0, T), \\ u(0) &= u_0, \end{aligned} \tag{45}$$

with the initial datum $u_0 \in \mathbb{R}$, the right hand side $f : (0, T) \to \mathbb{R}$ and some coefficient $\lambda \geq 0$. The model problem (45) is a simple test problem for the time-parallel multigrid method, which can be extended also to more general parabolic problems, see Gander and Neumüller (2016). Since these problems are usually stiff, implicit schemes should be used for the discretization of (45). To keep everything simple, we consider the implicit Euler scheme, i.e. we obtain

$$(1 + \lambda \Delta t) u_{k+1} = \Delta t \, f(t_{k+1}) + u_k \qquad \text{for } k = 0, 1, 2, \ldots, m, \tag{46}$$

where we use a uniform decomposition of the time interval $(0, T)$ into $m \in \mathbb{N}$ time steps of size $\Delta t = \frac{T}{m}$ and $t_k = k \Delta t$ for $k = 0, \ldots, m$. Each equation of (46) can be solved from one time step to the next time step, which is a purely sequential process. Here we put all the equations into one big linear system. We obtain

$$\mathbf{L} \underline{u} = \underline{f}, \tag{47}$$

with the system matrix

$$\mathbf{L} := \begin{pmatrix} 1 + \lambda \Delta t & & & & \\ -1 & 1 + \lambda \Delta t & & & \\ & \ddots & \ddots & & \\ & & -1 & 1 + \lambda \Delta t & \\ & & & -1 & 1 + \lambda \Delta t \end{pmatrix},$$

the coefficient vector and the right hand side

$$\underline{u} := \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_{m-1} \\ u_m \end{pmatrix} \quad \text{and} \quad \underline{f} := \begin{pmatrix} \Delta t \, f(t_1) + u_0 \\ \Delta t \, f(t_2) \\ \vdots \\ \Delta t \, f(t_{m-1}) \\ \Delta t \, f(t_m) \end{pmatrix}.$$

The idea now is to apply the multigrid method introduced in Sect. 6.3 to solve the linear system (47) at once. As in Sect. 6.1, we apply the damped Jacobi method, see Algorithm 5, to the linear system (47) for $\lambda = 1$, $T = 1$ and $\Delta t = 10^{-2}$. For the right hand side, we again use $\underline{f} = \underline{0}$. As in Sect. 6.1, we test the damped Jacobi scheme with a random initial guess $\underline{u}^{(0)} = [\text{rand}(0, 1)]_{j=1}^{m}$ for two different damping parameters $\tau \in \{1, \frac{1}{2}\}$. In Fig. 6, we plotted the error for different iterations. For the damping $\tau = 1$ the spectral radius of the iteration matrix of the Jacobi scheme is zero. So in that sense $\tau = 1$ would be the optimal damping parameter, since we obtain the exact solution in a finite number of Jacobi iterations, see also Fig. 6. We also observe, as in Sect. 6.1, that the optimal damping parameter does not reduce the high oscillatory

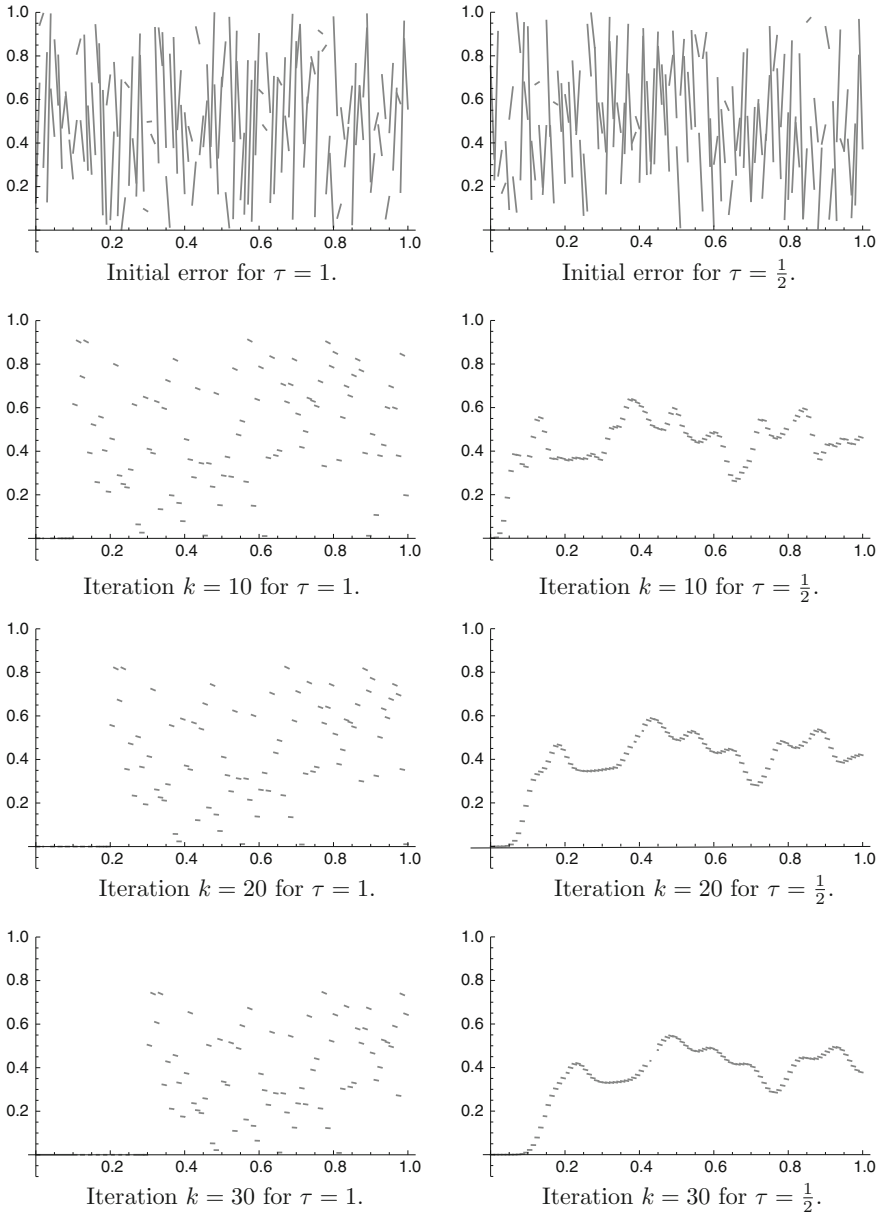


Fig. 6 Errors of the damped Jacobi scheme applied to the problem (47) for different iterations and two different damping parameters $\tau \in \{1, \frac{1}{2}\}$

part of the error. If we use, for example, $\tau = \frac{1}{2}$, we observe in Fig. 6 that the error gets smoother in each Jacobi iteration. This can be also verified by the so called Fourier mode analysis, see Gander and Neumüller (2016). Hence the best possible correction is again a smooth functions, which can usually be well approximated on a coarser time grid. Thus, we again can use the multigrid Algorithm 14 to solve the linear system (47). For this simple example the prolongation and restriction matrices are given by

$$\mathbf{P} = \begin{pmatrix} 1 & & \\ 1 & & \\ & 1 & \\ & 1 & \\ & & \ddots \\ & & 1 \\ & & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{R} = \mathbf{P}^\top.$$

For more advanced time stepping schemes, one can obtain the restriction and prolongation matrices by interpreting the time stepping scheme as a Galerkin scheme, see for example Gander and Neumüller (2016). We also observe that the damped Jacobi scheme can be applied in parallel for each time step. So we end up with a parallel method with respect to time. In the next example we demonstrate the parallel performance of this method

Example 6.2 Let us consider the example given in Gander and Neumüller (2016, Example 6.1). There the heat equation is solved on the three dimensional domain $\Omega = (0, 1)^3$ which is decomposed into 49 152 tetrahedra. For the time stepping scheme, a discontinuous Galerkin approach is used with polynomials of order three. This allows a time step size of $\Delta t = 10^{-1}$. The difference of this example to the simpler problem (47) is that now the diagonal entries are matrices instead of simple scalars. To apply now the damped Jacobi scheme, we have to invert these matrices. To do so, we further approximate the exact inverse of each of these matrices by applying one V-cycle of another multigrid cycle with respect to space. This results in a space-time multigrid approach which again has the advantage that it can be applied in parallel with respect to time. The parallel performance of this approach for this example is shown in the Tables 1 and 2. We observe perfect weak and strong scaling results. For more details see Gander and Neumüller (2016).

Cores	Time steps	dof	iter	Time	fwd. sub.
1	2	59 768	7	28.8	19.0
2	4	119 536	7	29.8	37.9
4	8	239 072	7	29.8	75.9
8	16	478 144	7	29.9	152.2
16	32	956 288	7	29.9	305.4
32	64	1 912 576	7	29.9	613.6
64	128	3 825 152	7	29.9	1 220.7
128	256	7 650 304	7	29.9	2 448.4
256	512	15 300 608	7	30.0	4 882.4
512	1 024	30 601 216	7	29.9	9 744.2
1 024	2 048	61 202 432	7	30.0	19 636.9
2 048	4 096	122 404 864	7	29.9	38 993.1
4 096	8 192	244 809 728	7	30.0	81 219.6
8 192	16 384	489 619 456	7	30.0	162 551.0
16 384	32 768	979 238 912	7	30.0	313 122.0
32 768	65 536	1 958 477 824	7	30.0	625 686.0
65 536	131 072	3 916 955 648	7	30.0	1 250 210.0
131 072	262 144	7 833 911 296	7	30.0	2 500 350.0
262 144	524 288	15 667 822 592	7	30.0	4 988 060.0

Table 1 Weak scaling

Cores	Time steps	dof	iter	Time
1	512	15 300 608	7	7 635.2
2	512	15 300 608	7	3 821.7
4	512	15 300 608	7	1 909.9
8	512	15 300 608	7	954.2
16	512	15 300 608	7	477.2
32	512	15 300 608	7	238.9
64	512	15 300 608	7	119.5
128	512	15 300 608	7	59.7
256	512	15 300 608	7	30.0
512	524 288	15 667 822 592	7	15 205.9
1 024	524 288	15 667 822 592	7	7 651.5
2 048	524 288	15 667 822 592	7	3 825.3
4 096	524 288	15 667 822 592	7	1 913.4
8 192	524 288	15 667 822 592	7	956.6
16 384	524 288	15 667 822 592	7	478.1
32 768	524 288	15 667 822 592	7	239.3
65 536	524 288	15 667 822 592	7	119.6
131 072	524 288	15 667 822 592	7	59.8
262 144	524 288	15 667 822 592	7	30.0

Table 2 Strong scaling

7 Concluding Remarks

In this chapter, we have introduced and discussed classical direct and iterative solvers. Direct methods lose efficiency for large-scale systems since the arithmetical work and memory demand are far from being optimal, whereas the classical iterative methods suffer from the bad conditioning of the systems arising form finite element discretization of PDEs and leading to large iteration numbers. In the former case, a smart elimination strategy can considerable improve the arithmetical complexity and the memory demand. In the latter case, preconditioning can help a lot.

In the framework of multigrid methods, direct methods and classical iterative methods enter into a perfect symbiosis. The damped Jacobi method or the Gauss–Seidel method are perfect smoothers whereas the systems arising in each multigrid iteration on the coarsest level are usually solved by some direct method. This symbiosis leads to multigrid methods that exhibit optimal complexity with respect to both arithmetical work and memory demand. The multigrid technique can be extended to time-dependent problems leading to time-parallel and space-time multigrid methods. A challenge for multigrid methods is the construction of a grid hierarchy when only one level of a mesh is given or even more, when only the system matrix itself is present. However it is possible to construct a hierarchy by only using the information about the system matrix. This leads to so-called algebraic multigrid methods, see for example Stüben (2001), Trottenberg et al. (2001) and references therein.

Of course, in this chapter, it is not feasible to discuss all possible classes of solvers for systems arising form the discretization of PDEs. We would like to mention only two further classes of solvers that are in particular relevant in Computational Acoustics. Domain decomposition methods (DDMs) deliver the technique to construct solvers that are highly suited for implementation on massively parallel computers. We refer the reader to the monographs by Douglas et al. (2003), Toselli and Widlund (2005), Pechstein (2013) and Korneev and Langer (2015) for an introduction to different domain decomposition methods. The literature on DDMs is now very rich. In particular, the proceedings of the International Conferences on DDMs,

that can be found on the DD website,[2] provide an overview of the development of DDMs. System matrices arising from the discretization by means of the Boundary Element Method (BEM) are smaller than the corresponding finite element stiffness matrices, but the BEM matrices are fully populated. Thus, data sparse techniques for the approximate representation of these matrices and a corresponding calculus have been developed, see, e.g., Rjasanow and Steinbach (2007), Steinbach (2008), Hackbusch (2009), and the references therein.

References

Bakhvalov, N. S. (1966). On the convergence of a relaxation method with natural constraints on the elliptic operator. *USSR Computational Mathematics and Mathematical Physics*, *6*(5), 101–135.

Braess, D., & Hackbusch, W. (1983). A new convergence proof for the multigrid method including the V-cycle. *SIAM Journal on Numerical Analysis*, *20*(5), 967–975.

Bramble, J. (1993). *Multigrid methods* (Vol. 294). Pitman research notes in mathematical sciences. Harlow: Longman Scientific and Technical.

Bramble, J. H., Pasciak, J. E., & Xu, J. (1990). Parallel multilevel preconditioners. *Mathematics of Computation*, *55*, 1–22.

Davis, T. A. (2006). *Direct methods for sparse linear systems*. Philadelphia: SIAM.

Douglas, C. C., Haase, G., & Langer, U. (2003). *A tutorial on elliptic PDE solvers and their parallelization*. Software, environments, and tools. Philadelphia: SIAM.

Duff, I. S., Erisman, A. M., & Reid, J. K. (1986). *Direct methods for sparse matrices*. Oxford: Oxford University Press.

Fedorenko, R. P. (1961). A relaxation method for solving elliptic difference equations. *USSR Computational Mathematics and Mathematical Physics*, *1*(5), 1092–1096.

Fedorenko, R. P. (1964). The speed of convergence of one iterative process. *USSR Computational Mathematics and Mathematical Physics*, *4*(3), 227–235.

Gander, M. J., & Neumüller, M. (2016). Analysis of a new space-time parallel multigrid algorithm for parabolic problems. *SIAM Journal on Scientific Computing*, *38*(4), A2173–A2208.

George, A., & Liu, J. W. H. (1981). *Computer solutions of large sparse positive definite systems*. Englewood Cliffs: Prentice Hall.

Hackbusch, W. (1985). *Multi-grid methods and applications*. Berlin: Springer.

Hackbusch, W. (2009). *Hierarchische Matrizen: Algorithmen und Analysis*. Berlin: Springer.

Hestenes, M. R., & Stiefel, E. (1952). Methods of conjugate gradients for solving linear systems. *Journal of Research of the National Bureau of Standards*, *49*(6), 409–436.

Jung, M., & Langer, U. (2013). *Methode der finiten Elemente für Ingenieure: Eine Einführung in die numerischen Grundlagen und Computersimulation* (2nd ed.). Berlin: Springer.

Korneev, V. G., & Langer, U. (2015). *Dirichlet-Dirichlet domain decomposition methods for elliptic problems: h and hp finite element discretizations*. New Jersey: World Scientific Publishing Company Incorporated.

Oswald, P. (1999). On the robustness of the BPX-preconditioner with respect to jumps in the coeffcients. *Mathematics of Computation*, *68*(226), 633–650.

Pechstein, C. (2013). *Finite and boundary element tearing and interconnecting solvers for multiscale problems* (Vol. 90). Lecture notes in computational science and engineering. Berlin: Springer.

Rjasanow, S., & Steinbach, O. (2007). *The fast solution of boundary integral equations*. Mathematical and analytical techniques with applications to engineering. Berlin: Springer.

Saad, Y. (2003). *Iterative methods for sparse linear systems* (2nd ed.). Philadelphia: SIAM.

[2] http://www.ddm.org/conferences.html.

Steinbach, O. (2008). *Numerical approximation methods for elliptic boundary value problems: Finite and boundary elements*. Berlin: Springer.
Stüben, K. (2001). A review of algebraic multigrid. *Journal of Computational and Applied Mathematics*, *128*. Numerical analysis 2000 (Vol. VII). Partial differential equations.
Toselli, A., & Widlund, O. (2005). *Domain decomposition methods - algorithms and theory* (Vol. 34). Springer series in computational mathematics. Berlin: Springer.
Trottenberg, U., Oosterlee, C. W., & Schüller, A. (2001). *Multigrid*. San Diego: Academic Press Inc.
Zhang, X. (1992). Multilevel Schwarz methods. *Numerische Mathematik*, *63*(4), 521–539.
Zlatev, Z. (1991). *Computational methods for general sparse matrices*. Dordrecht: Kluwer Academic Publishers.