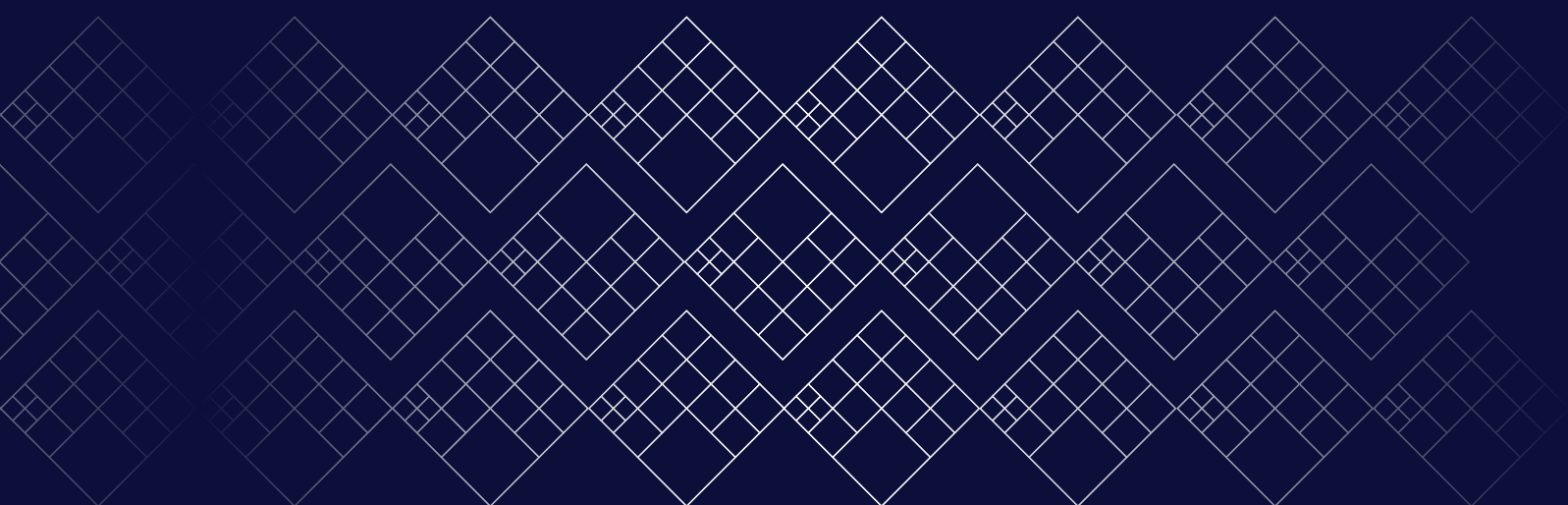
Olena Palii

On approximation methods and efficient iterative solvers for the Radiative Transfer Equation



On approximation methods and efficient iterative solvers for the Radiative Transfer Equation

Olena Palii

ON APPROXIMATION METHODS AND EFFICIENT ITERATIVE SOLVERS FOR THE RADIATIVE TRANSFER EQUATION

Dissertation

to obtain
the degree of doctor at the University of Twente,
on the authority of the rector magnificus,
prof.dr.ir. A. Veldkamp,
on account of the decision of the Doctorate Board,
to be publicly defended
on Thursday 2 of June 2022 at 14.45 hours

by

Olena Palii

born on 26 of February 1994
in Kyiv, Ukraine

This dissertation has been approved by:
Supervisor:
prof.dr.ir. J.J.W. van der Vegt
Co-supervisor:
dr. M. Schlottbom

Cover design: The cover was designed by Daria Fonina.
Printed by: Gildeprint
ISBN: 978-90-365-5388-9
DOI: 10.3990/1.9789036553889

GRADUATION COMMITTEE:

Chairman/secretary	prof.dr. J.N. Kok (University of Twente)
Supervisor	prof.dr.ir. J.J.W. van der Vegt (University of Twente)
Co-supervisor	dr. M. Schlottbom (University of Twente)
Members	prof.dr.ir. E.H. van Brummelen (TU Eindhoven)
	prof.dr. C. Brune (University of Twente)
	prof.dr. H. Egger (Johannes Kepler University Linz)
	prof.dr. M. Frank (Karlsruher Institute of Technology)
	prof.dr. P.J. Kelly (University of Twente)

This work was carried out at the
Mathematics of Computational Science (MACS) chair,
Faculty of Electrical Engineering, Mathematics and Computer Science,
University of Twente, P.O. Box 217,
7500 AE Enschede, The Netherlands.

To my mother and grandmother

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Studies about light and applications of radiative transfer models

When in need to model the transport of photons and their interaction with a medium consisting of randomly distributed particles, the radiative transfer equation (RTE), also called Linear Boltzmann equation, is a popular choice [30]. The RTE is an integro-differential equation that describes the distribution of photons in position and velocity space in a medium. A detailed description of the RTE will be provided in Section 1.2.

The RTE allows to calculate the specific intensity of light while it is propagating through media with different material characteristics. The RT model takes into account properties of the photon flux and media, such as absorption, emission and scattering, thus allowing for a wide range of applications, as outlined next.

In atmospheric science the use of the RTE allows the description of radiation passing through the atmosphere of the Earth, which is modelled with the so called plane-parallel model. One application of the plane-parallel model is satellite data assimilation [21]. Another example is a so-called free space optical communication system, which is used for transmitting telecommunication data wirelessly [3].

A prominent example in remote sensing is the use of the RTE for optical oceanography purposes. Here measurements collected remotely by aircraft or satellite are used to recover data about the properties and condition of oceans and seas, the physical characteristics of the bottom, etc. [4]. There are also interesting applications of the RTE in the field of cloud tomography, like the retrieval of cloud optical thickness [31]. This helps to better understand the function of clouds in the climate system of the Earth.

In stellar spectrometry, due to the scale of the problem, a plane-parallel model provides a useful simplification of the complex physics, for instance see [5], in which a plane-parallel geometry is considered to study the structure of galaxies.

In biomedical studies, the RTE is a suitable choice for modeling light propagation in biological tissues, see for instance [6, 7, 8, 9, 10]. In biomedical studies the RTE is generally used as a forward model in related inverse problems, which aim to recover biomedical information from data at the boundary [13].

1.2 The radiative transfer boundary-value problem in multiple dimensions

We consider the steady-state radiative transfer equation. The steady-state RTE is applicable for cases where the propagation of photons is happening much faster than other physical processes. An example of an application where the usage of the steady-state equation is justified can be found in photo-acoustic imaging applications [13], as sound waves propagate much slower than light in biomedical tissue.

The RTE is formulated in multiple dimensions as

$$s \cdot \nabla_r \phi(r,s) + \sigma_t(r)\phi(r,s) = q(r,s) + \sigma_s(r) \int_{S^2} k(s \cdot s')\phi(r,s')ds', \qquad (1.1)$$

and is complemented by the boundary condition

$$\phi(r,s) = g(r,s) \text{ for } (r,s) \in \partial R \times S, \text{ such that } s \cdot n(r) < 0. \qquad (1.2)$$

Here, $\phi(r,s)$ is a density function representing, for example, the density of photons. The density function depends on the position $r \in R \subseteq \mathbb{R}^3$, and the direction of propagation $s \in S^2$. In general the velocity domain S^2 is the unit sphere and $s \in S^2$ is a vector variable that describes the direction of propagation. Furthermore, we denote with $n(r)$ the unit outer normal vector on the boundary ∂R.

On the left-hand side of (1.1) the transport of photons is modelled by a directional derivative $s \cdot \nabla_r \phi(r,s)$. The second term $\sigma_t(r)\phi(r,s)$ in (1.1) is one of the terms that describe the interaction of the quantity ϕ with the medium. The coefficient function $\sigma_t(r) = \sigma_a(r) + \sigma_s(r)$ is called the total attenuation cross section, and $\sigma_a(r)$ and $\sigma_s(r)$ are, respectively, the absorption and scattering coefficients.

On the right-hand-side, the scattering of photons is modelled by the integral operator $\sigma_s(r) \int_{S^2} k(s \cdot s')\phi(r,s')ds'$, and $q(r,s)$ represents the source function.

1.3 Slab geometry model

In this section we will derive a slab geometry radiative transfer equation from the original 3D problem (1.1)-(1.2). We assume translational invariance, that is the physical domain is infinite in two directions, i.e. $\Omega = \mathbb{R}^2 \times (0, L)$, and the photon density is independent of (x, y) and the azimuthal angle ψ. Moreover, we assume that the optical parameters and data depend only on z and polar angle θ. A schematic drawing of the one-layer slab geometry domain is given on the left hand side of Figure 1.1. Using spherical coordinates, the vector s can be expressed using the polar (θ) and azimuthal (ψ) angles

$$s = (s_1, s_2, s_3) = (\cos\psi \sin\theta, \sin\psi \sin\theta, \cos\theta)^T.$$

We next derive expressions for the transport, attenuation, and source terms for the slab geometry. Since in a slab geometry ϕ does not depend on x, y and ψ, the term $s \cdot \nabla\phi$ reduces to $\mu \frac{\partial \phi}{\partial z}$ with $\mu = \cos\theta$,

$$s \cdot \nabla\phi = s_1\partial_x\phi + s_2\partial_y\phi + s_3\partial_z\phi = \cos\theta \frac{\partial\phi}{\partial z} = \mu \frac{\partial\phi}{\partial z}.$$

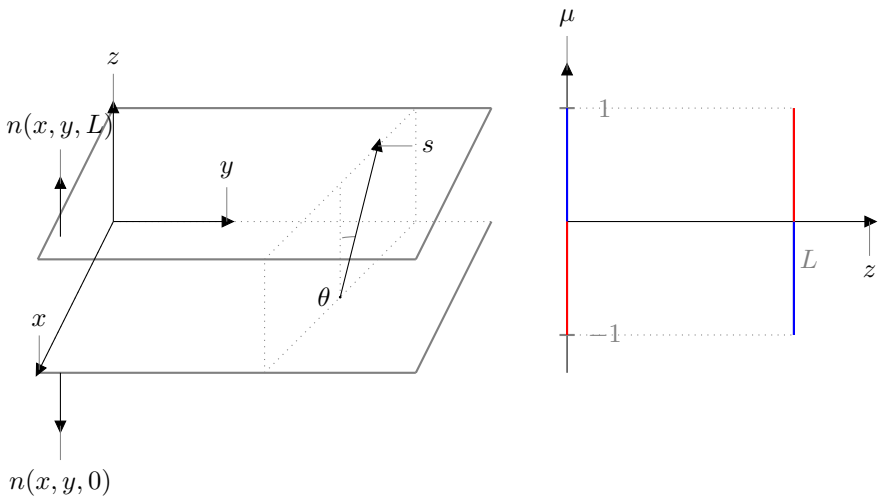


Figure 1.1: Left: Slab geometry in physical coordinates. Right: Slab in (z, μ) plane, with inflow boundary in blue, and outflow boundary in red.

The scattering kernel $k(s \cdot s')$ will reduce in the slab geometry model to $\tilde{k}(\mu, \mu')$, as explained next. Consider

$$s \cdot s' = \sin\theta \sin\theta' \cos(\psi - \psi') + \cos\theta \cos\theta'.$$

Then by denoting $\mu = \cos\theta$ and $\mu' = \cos\theta'$, we get

$$s \cdot s' = (1-\mu^2)^{1/2}(1-(\mu')^2)^{1/2} \cos(\psi - \psi') + \mu\mu'.$$

After substituting $\psi - \psi' = \bar{\psi}$ and integrating the scattering kernel over the sphere, we obtain that

$$\begin{aligned}
&\int_{S^2} k(s \cdot s')\phi(z, s'_3) ds' = \\
&\int_0^{\pi} \int_0^{2\pi} k(\sin\theta \sin\theta' \cos\bar{\psi} + \cos\theta\cos\theta')\phi(z, \cos\theta') \sin\theta' d\bar{\psi} d\theta' = \\
&\int_{-1}^{1} \left(\int_0^{2\pi} k((1-\mu^2)^{1/2}(1-(\mu')^2)^{1/2} \cos\bar{\psi} + \mu\mu') d\bar{\psi} \right) \phi(z, \mu') d\mu',
\end{aligned}$$

where we used the substitution $\mu = \cos\theta$ and $\mu' = \cos\theta'$. This motivates the definition

$$\tilde{k}(\mu, \mu') = \int_0^{2\pi} k((1-\mu^2)^{1/2}(1-(\mu')^2)^{1/2} \cos\bar{\psi} + \mu\mu') d\bar{\psi}.$$

Then the slab geometry radiative transfer problem can be written as

$$\mu \frac{\partial \phi(z,\mu)}{\partial z} + \sigma_t(z)\phi(z,\mu) = \sigma_s(z) \int_{-1}^{1} \tilde{k}(\mu, \mu')\phi(z, \mu') d\mu' + q(z, \mu). \tag{1.3}$$

In full analogy with the multidimensional case, equation (1.3) requires boundary conditions to be defined. Consider the normal vector $n(x, y, 0) = (0, 0, -1)$,

see Figure 1.1. Using the expression for s in spherical coordinates we obtain $s \cdot n = -\cos\theta = -\mu$, hence the boundary condition at $z = 0$ is defined for $\mu > 0$. Similarly for the outward normal vector $n(x, y, L) = (0, 0, 1)$, $s \cdot n = \cos\theta = \mu$ and the boundary condition at $z = L$ is defined for $\mu < 0$. Summarizing, the boundary condition is prescribed by

$$\begin{cases} \phi(0,\mu) & = g^0(\mu), \text{ if } \mu > 0, \\ \phi(L,\mu) & = g^L(\mu), \text{ if } \mu < 0. \end{cases} \tag{1.4}$$

Multilayered media

The slab geometry RTE can also be used to model a beam of light that propagates through a composite material with a multilayered structure. In this case we consider a change in the refractive index when transitioning from one layer to another. This introduces reflection and transmission conditions at the corresponding interfaces. If the refractive index does not change across layers then a multilayered slab is included in (1.3) - (1.4), when piecewise constant spectral coefficients are used. Consider a schematic description of a multilayered slab, see Figure 1.2.

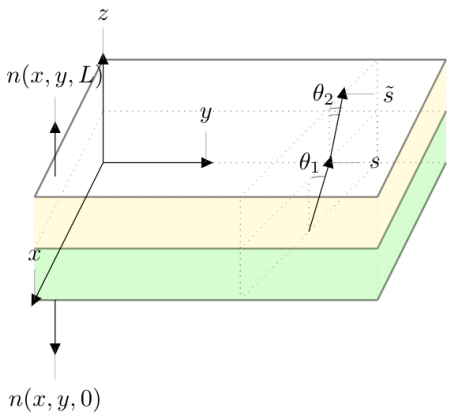


Figure 1.2: Layered slab geometry in physical coordinates, layer 1(green), layer 2(yellow).

By considering the RTE in layered media we can model a situation, in which photons propagate through media with varying spectral properties. There are a number of applications for which such layered media are important.

In [25], the estimation of the albedo and optical thickness parameters from the collected measurements in a stratified atmosphere is considered. The atmosphere is essentially translated into a two-layered medium in slab geometry. In remote sensing applications [28] a layered slab geometry allows to model the layers of the atmosphere or different layers of the ocean, such as surface, the bottom covering layer of vegetation and soil, and to retrieve their optical characteristics.

A good classification of radiative transfer inverse problems in multilayered slab geometry is given in [29], and more solution methods can be found in [27, 29, 26].

1.4 Overview of classical approximation techniques

In this section we give an overview of several classical methods that are used to discretize and solve radiative transfer problems numerically. For ease of presentation, we will discuss these techniques for the slab geometry RTE problem (1.3)-(1.4).

Legendre polynomial expansion

A popular technique for approximating the angular component of the solution ϕ in (1.3)-(1.4) uses a truncated series of Legendre polynomials $\{P_l\}_l$, which is known as the P_N approximation (or spherical harmonics expansion in higher dimensions)[22, 23],

$$\phi(z,\mu) = \sum_{l=0}^{\infty} \frac{2l+1}{2} \phi_l(z) P_l(\mu). \tag{1.5}$$

Legendre polynomials are orthogonal on $[-1,1]$ and satisfy the following orthogonality relation

$$\int_{-1}^{1} P_l(\mu) P_{l'}(\mu) d\mu = \frac{2}{2l+1} \delta_{ll'}.$$

For these properties please refer to [2]. The scattering term in (1.3) can also be represented using Legendre polynomials, which in general would require the use of the addition theorem for spherical harmonics. Consider, for simplicity, the isotropic scattering case, where $\tilde{k}(\mu,\mu') = \frac{1}{2}$. Then, using the orthogonality of the Legendre polynomials, we obtain

$$\frac{1}{2}\int_{-1}^{1} \phi(z,\mu')d\mu' = \frac{1}{2}\sum_{l=0}^{\infty} \frac{2l+1}{2}\phi_l(z) \int_{-1}^{1} P_l(\mu')d\mu' = \phi_0(z).$$

Then, after substituting expression (1.5) into (1.3), assuming that $\phi_{L+1} = 0$ and taking the inner product with $P_{l'}$ for $l' = 0, \ldots, L$ in the angular domain we obtain

$$\begin{aligned} &\sum_{l=0}^{L} \phi_l'(z) \int_{-1}^{1} \frac{2l+1}{2} \mu P_l(\mu) P_{l'}(\mu) d\mu + \sigma_t(z)\phi_{l'}(z) = \\ &\frac{\sigma_s(z)}{2}\phi_0(z)\delta_{0,l'} + \int_{-1}^{1} q(z,\mu) P_{l'}(\mu) d\mu, \ l' = 0, \ldots, L, \end{aligned} \tag{1.6}$$

where $\phi_l'(z) = \frac{d\phi_l(z)}{dz}$. Using the recurrence relation for Legendre polynomials,

$$(2l+1)\mu P_l(\mu) = (l+1)P_{l+1}(\mu) + lP_{l-1}(\mu),$$

we can write (1.6) as a coupled system of $L+1$ equations with $L+1$ unknowns ϕ_l, for $l = 0, \ldots, L$, which, by setting $\phi_{-1} = 0$ and $\phi_{L+1} = 0$, is given by

$$\frac{l+1}{2l+1}\phi'_{l+1}(z) + \frac{l}{2l+1}\phi'_{l-1}(z) + \sigma_t(z)\phi_l(z) = \frac{\sigma_s(z)}{2}\phi_0(z)\delta_{0,l} + \int_{-1}^{1} q(z,\mu)P_l(\mu)d\mu.$$

We also must impose $L+1$ boundary conditions on $\phi_l(z)$, $l = 0, .., L$. Formulating boundary conditions in the P_N method is a subject of discussion, because the representation of the solution via finite set of Legendre polynomials is in general inconsistent with the boundary conditions, see [24]. One of the possibilities is to use the approximate formulation of the boundary condition, which results from taking the inner product with the Legendre polynomials

$$\begin{cases} \sum_{l=0}^{L} \int_0^1 \frac{2l+1}{2}\phi_l(0)P_l(\mu)P_{l'}(\mu)d\mu & = \int_0^1 g^0(\mu)P_{l'}(\mu)d\mu, \\ \sum_{l=0}^{L} \int_{-1}^0 \frac{2l+1}{2}\phi_l(L)P_l(\mu)P_{l'}(\mu)d\mu & = \int_{-1}^0 g^L(\mu)P_{l'}(\mu)d\mu. \end{cases}$$

Since Legendre polynomials are not orthogonal in $L^2(-1,0)$ and $L^2(0,1)$ such an approximation of the boundary condition results in a dense coupling of the functions ϕ_l, see [32] for more details.

Discrete ordinates method

Another popular discretization method for the slab geometry model (1.3)-(1.4) that is used to describe the angular dependence is the discrete ordinates method or S_N method, introduced in [1]. The idea is based on approximating the infinite angular space with a finite set of N directions and can be viewed as a collocation method.

Consider the slab geometry problem (1.3)-(1.4) with isotropic scattering term $\tilde{k}(\mu,\mu') = \frac{1}{2}$, namely

$$\mu\frac{\partial\phi(z,\mu)}{\partial z} + \sigma_t(z)\phi(z,\mu) = \frac{\sigma_s(z)}{2}\int_{-1}^{1}\phi(z,\mu')d\mu' + q(z,\mu),$$

for $z \in (0, L)$ and $\mu \in (-1, 1)$, with the inflow boundary condition

$$\begin{cases} \phi(0,\mu) & = g^0(\mu), \text{ if } \mu > 0, \\ \phi(L,\mu) & = g^L(\mu), \text{ if } \mu < 0. \end{cases}$$

Let the set of discrete ordinates consist of the quadrature points μ_i, such that $\mu_i \neq 0$, and corresponding weights ω_i for $i = 1, .., N$, where N is the chosen number of directions of propagation. Let $\phi_i(z) \approx \phi(z,\mu_i)$ be defined by

$$\mu_i\phi'_i(z) + \sigma_t(z)\phi_i(z) = \frac{\sigma_s(z)}{2}\sum_{j=1}^{N}\phi_j(z)\omega_j + q(z,\mu_i), \tag{1.7}$$

where $\phi'_i(z) = \frac{d\phi_i(z)}{dz}$, and

$$\begin{cases} \phi_i(0) & = g^0(\mu_i), \text{ if } \mu_i > 0, \\ \phi_i(L) & = g^L(\mu_i), \text{ if } \mu_i < 0. \end{cases} \tag{1.8}$$

In this way the radiative transfer equation is represented as a system of advection equations with corresponding boundary conditions for each chosen direction of propagation that is coupled via the sum on the right-hand side of (1.7).

The sum in (1.7) requires special attention when solving for ϕ_l, $l = 1, .., N$. If the summation in (1.7) can be performed efficiently, solving (1.7)–(1.8) iteratively is a viable approach.

Iterative solvers

The system of equations (1.7)–(1.8) resulting from the discretization of the RTE using the discrete ordinates method can be solved iteratively, for example with the Source Iteration (SI) method [16], which will be explained in more detail in Chapter 2. This contraction mapping proved to be mostly effective for optically thin regimes and very slow for diffusive regimes [16], where $c = ||\sigma_s/\sigma_t||_\infty \approx 1$. In order to address these problems acceleration techniques, i.e., preconditioners were proposed. The Diffusion Synthetic Acceleration (DSA) [14] can be used for isotropic scattering, but for strongly forwardly peaked scattering the convergence rate deteriorates [16, 17, 18]. An angular multigrid preconditioner for the slab geometry was proposed in [18] as an alternative to DSA, and the convergence analysis is done in [19]. In multiple dimensions the multigrid preconditioner developed for the source iteration does not converge, but can be effective when implemented as a preconditioner for Krylov-based techniques, see [15]. Several multigrid preconditioners were developed for various iterative techniques, see [11, 12, 20]. However, the construction of robustly convergent schemes for the most general cases is still an open problem.

1.5 Thesis outline

To conclude the introduction we formulate the main research questions, that will be addressed in this thesis.

1. Developing a robust preconditioner for the source iteration technique. For ease of implementation and presentation this will be considered first for the case of isotropic scattering.

2. Extending the developed preconditioners to anisotropic scattering problems.

3. Applying the discretized version of the scattering operator efficiently to solve the anisotropic scattering problem robustly.

4. Solving isotropic scattering problems with a stable and accurate non-tensor product phase-space discretization technique suitable for adaptive mesh refinement.

The outline of the thesis is as follows: in Chapter 2 we will give a more detailed description of the source iteration method, and relate it to another common solution technique, namely the first collision source method. In Chapter 3 we will propose a generalisation of the S_N and P_N methods in combination with a robust preconditioned source iteration method. This method is implemented

for isotropic radiative transfer problems, addressing research question 1. In Chapter 4 we will discuss the efficient solution of the linear equations that arise from the discretization of the anisotropic radiative transfer. We will propose a robust and efficient iterative method for solving this problem as well as an efficient application of the anisotropic scattering operator, addressing research questions 2 and 3. In Chapter 5 we address research question 4, and we provide an analysis of a discontinuous Galerkin approximation in both space and angle that does not use a tensor product formulation, allowing for adaptive discretization in phase-space. This algorithm is implemented for the isotropic radiative transfer equation in slab geometry for problems with non-smooth and discontinuous solutions.

Bibliography

[1] S. Chandrasekhar: Radiative Transfer. Dover Publications, Inc. (1960)

[2] M. Case, P. F. Zweifel: Linear transport theory. Addison-Wesley, Reading (1967)

[3] S.S. Muhammad, M.Z. Malik, G. Kandus: Optical propagation modelling using radiative transfer equation (RTE). 16th International Conference on Transparent Optical Networks (ICTON), 1–4, IEEE (2014)

[4] C.D. Mobley: Radiative transfer in the ocean. Encyclopedia of ocean sciences, 2321-2330 (2001)

[5] M. Baes, H. Dejonghe: Radiative transfer in disc galaxies–I. A comparison of four methods to solve the transfer equation in plane-parallel geometry. Monthly Notices of the Royal Astronomical Society 326, 722–732 (2001)

[6] A.D. Klose, A. Bluestone, M. Löcker, G. Abdoulaev, A.H. Hielscher, J. Beuthan: Optical tomography with the equation of radiative transfer. Biomedical Optical Spectroscopy and Diagnostics, Optical Society of America (2000)

[7] A.D. Klose, U. Netz, J. Beuthan, A.H. Hielscher: Optical tomography using the time-independent equation of radiative transfer—Part 1: forward model. Elsevier Journal of Quantitative Spectroscopy and Radiative Transfer 72(5), 691–713 (2002)

[8] A.D. Klose, A.H. Hielscher: Optical tomography using the time-independent equation of radiative transfer—Part 2: inverse model. Elsevier Journal of Quantitative Spectroscopy and Radiative Transfer 72(5), 715–732 (2002)

[9] K. Ren, G. Bal, A.H. Hielscher: Frequency domain optical tomography based on the equation of radiative transfer. SIAM Journal on Scientific Computing 28(4), 1463-1489 (2006)

[10] K. Ren, B. Moa-Anderson, G. Bal, X. Gu, A.H. Hielscher: Frequency domain tomography in small animals with the equation of radiative transfer. Optical Tomography and Spectroscopy of Tissue VI, International Society for Optics and Photonics 5693, 111-120 (2005)

[11] P.N. Brown, B. Lee, T.A. Manteuffel: A moment-parity multigrid preconditioner for the first-order system least-squares formulation of the Boltzmann transport equation. SIAM Journal on Scientific Computing 25(2), 513–533 (2003)

[12] G. Kanschat, J.-C. Ragusa: A robust multigrid preconditioner for S_NDG approximation of monochromatic, isotropic radiation transport problems. SIAM Journal on Scientific Computing 36(5), A2326-A2345 (2014)

[13] S.R. Arridge: Optical tomography in medical imaging. Inverse problems. An International Journal on the Theory and Practice of Inverse Problems, Inverse Methods and Computerized Inversion of Data 15(2), R41–R93 (1999)

[14] B. Turcksin, J.C. Ragusa: Discontinuous diffusion synthetic acceleration for S_N transport on 2D arbitrary polygonal meshes. Elsevier Journal of Computational Physics 274, 356–369 (2014)

[15] B. Turcksin, J.C. Ragusa, J.E. Morel: Angular multigrid preconditioner for Krylov-based solution techniques applied to the S_N equations with highly forward-peaked scattering. Transport Theory and Statistical Physics, Taylor & Francis 41(1-2), 1-22 (2012)

[16] M.L. Adams, E.W. Larsen: Fast iterative methods for discrete-ordinates particle transport calculations. Elsevier, Progress in nuclear energy 40(1), 3-159 (2002)

[17] B. Lee: A novel multigrid method for S_n discretizations of the monoenergetic Boltzmann transport equation in the optically thick and thin regimes with anisotropic scattering, Part I. SIAM Journal on Scientific Computing 31(6), 4744-4773 (2010)

[18] J.E. Morel, T.A. Manteuffel: An angular multigrid acceleration technique for S_n equations with highly forward-peaked scattering. Taylor & Francis, Nuclear Science and Engineering 107(4), 330-342 (1991)

[19] S. Oliveira: Analysis of a multigrid method for a transport equation by numerical Fourier analysis. Elsevier, Computers & Mathematics with Applications 35(12), 7-12 (1998)

[20] D. Lathouwers, Z. Perkó: An angular multigrid preconditioner for the radiation transport equation with Fokker-Planck scattering. Elsevier Journal of Computational and Applied Mathematics 350, 165-177 (2019)

[21] J. Vidot: Overview of the status of radiative transfer models for satellite data assimilation. Seminar on Use of Satellite Observations in Numerical Weather Prediction, 8-12 (2014)

[22] G.C. Pomraning, M. Clark Jr.: The variational method applied to the monoenergetic Boltzmann equation. Part II. Nuclear Science and Engineering, 16(2), 155-164 (1963)

[23] M.F. Modest: Further development of the elliptic PDE formulation of the P_N approximation and its Marshak boundary conditions. Numerical Heat Transfer, Part B: Fundamentals, 62(2-3), 181-202 (2012)

[24] R.E. Marshak: Note on the spherical harmonic method as applied to the Milne problem for a sphere. Physical Review, 71(7), 443-446 (1947)

[25] R.E. Bellman, H.H. Kagiwada, R.E. Kalaba, S. Ueno: Inverse problems in radiative transfer: Layered media. Icarus, 4(2), 119-126 (1965)

[26] A. Da Silva, M. Elias, C. Andraud, J. Lafait: Comparison of the auxiliary function method and the discrete-ordinate method for solving the radiative transfer equation for light scattering. JOSA A, 20(12), 2321-2329 (2003)

[27] M. Machida, G.Y. Panasyuk, J.C. Schotland, V.A. Markel: The Green's function for the radiative transport equation in the slab geometry. Journal of Physics A: Mathematical and Theoretical, 43(6), 065402 (2010).

[28] R. Quast, W. Wagner: Analytical solution for first-order scattering in bistatic radiative transfer interaction problems of layered media. Applied optics, 55(20), 5379-5386 (2016).

[29] N.J. McCormick: Inverse radiative transfer problems: a review. Nuclear science and Engineering, 112(3), 185-198 (1992).

[30] M.I. Mishchenko: Poynting–Stokes tensor and radiative transfer in discrete random media: the microphysical paradigm. Optics express, 18(19), 19770-19791 (2010)

[31] T. Nakajima, M.D. King: Determination of the optical thickness and effective particle radius of clouds from reflected solar radiation measurements. Part I: Theory. Journal of Atmospheric Sciences, 47(15), 1878-1893 (1990).

[32] H. Egger, M. Schlottbom: A Perfectly Matched Layer Approach for P_N-Approximations in Radiative Transfer. SIAM Journal on Numerical Analysis, 57(5), 2166-2188 (2019).

Chapter 2

On the equivalence of the source iteration method and the first collision source method

2.1 Introduction

Iterative methods to solve the RTE [3, 4, 5] are a viable alternative to direct methods. Two of the most comprehensive studies on the source iteration method and supportive acceleration methods can be found in [2, 6]. In this chapter we will provide a summary of the source iteration method and an equivalent method - the first collision source method, which are often used to solve the radiative transfer equation. While the source iteration method and the first collision source method are equivalent, there are differences in how the solution can be obtained using each method.

Various approximation methods were developed and successfully applied for different media and geometries, for a brief overview see Section 1.4.

2.2 Source Iteration method

Recall the radiative transfer equations (1.1)–(1.2)

$$\begin{aligned} s \cdot \nabla_r \phi(r,s) + \sigma_t(r)\phi(r,s) &= q(r,s) + \\ &+ \sigma_s(r) \int_{S^2} k(s \cdot s')\phi(r,s')ds', \text{ for } (r,s) \in R \times S^2, \\ \phi(r,s) &= g(r,s), \text{ for } (r,s) \in \Gamma_-, \end{aligned}$$

where $R \subseteq \mathbb{R}^3$. Here the inflow boundary Γ_- is defined as

$$\Gamma_- = \{(r,s) \in \partial R \times S, s \cdot n(r) < 0\}, \tag{2.1}$$

see Section 1.2.

The basic idea behind the source iteration method lies in the following. We start with a given initial iterate ϕ_{k-1} and compute ϕ_k by solving the decoupled transport equations,

$$\begin{aligned} s \cdot \nabla_r \phi_k(r,s) + \sigma_t(r)\phi_k(r,s) &= \sigma_s(r) K \phi_{k-1}(r,s) + \qquad (2.2) \\ &+ q(r,s) \text{ for } (r,s) \in R \times S^2, \\ \phi_k(r,s) &= g(r,s) \text{ for } (r,s) \in \Gamma_-. \qquad (2.3) \end{aligned}$$

This process continues until the iteration converges within a specified tolerance.

The source iteration method is a well-researched method, and several results of its convergence can be found in the literature. Let

$$\|\phi\|_{\sigma_t}^2 = (\sigma_t\phi, \phi) = \int_R \int_{S^2} \sigma_t |\phi|^2 ds dr.$$

The following sources provide the proofs of convergence of the source iteration method

- Case and Zweifel [7, p.291]: $\|\phi - \phi_k\|_{\sigma_t} \leq \|\frac{\sigma_s}{\sigma_t}\|_\infty \|\phi - \phi_{k-1}\|_{\sigma_t}$, when $||\frac{\sigma_s}{\sigma_t}||_\infty < 1$.
- Blake [8, sec. 2.5.2 - 2.5.3], slab geometry: for the isotropic scattering case with $||\frac{\sigma_s}{\sigma_t}||_\infty \leq 1$.
- Egger and Schlottbom [1], for the case of bounded domains with $||\frac{\sigma_s}{\sigma_t}||_\infty \leq 1$.

2.3 First Collision Source Method

The First Collision Source Method (FCSM) was presented in Lathrop [9], and later also in Alcouffe [10]. It was also extended and modified for unstructured grids in multiple dimensions in [12] and implemented with adaptive quadrature in 3D [13].

The first collision source method decomposes the intensity into two components,

$$\phi(r, s) = \phi_c(r, s) + \phi_u(r, s),$$

where the physical interpretation of ϕ_c and ϕ_u is as follows:

- ϕ_u is the uncollided part, which corresponds to particles that have not undergone scattering events.
- ϕ_c is the collided component, which corresponds to the scattered particles.

By linearity of (1.1) and (1.2), we can define ϕ_u and ϕ_c as the solutions to

$$s \cdot \nabla_r \phi_u(r, s) + \sigma_t(r)\phi_u(r, s) = q(r, s), \tag{2.4}$$
$$s \cdot \nabla_r \phi_c(r, s) + \sigma_t(r)\phi_c(r, s) = \sigma_s(r)K\phi_c(r, s) + \sigma_s(r)K\phi_u(r, s), \tag{2.5}$$

for $(r, s) \in R \times S^2$, with corresponding boundary conditions

$$\phi_u(r, s) = g(r, s) \text{ for } (r, s) \in \Gamma_-, \tag{2.6}$$
$$\phi_c(r, s) = 0 \text{ for } (r, s) \in \Gamma_-. \tag{2.7}$$

Each component can be computed separately:

- ϕ_u solves an advection equation without scattering and can be computed numerically or analytically by using, for instance, the discrete ordinates method;
- ϕ_c cannot be computed analytically due to the inclusion of the collision operator, i.e., the computation of ϕ_c from (2.5), (2.7) is as difficult as the computation of ϕ itself.

The biggest potential advantage of FCSM is that the equation for the uncollided and collided parts can be solved using different approximation techniques. This was done in [10], where the problem (2.4), (2.6) is solved by the Monte-Carlo method and (2.5), (2.7) is solved using the discrete ordinates method.

Since the collided component ϕ_c is expected to have less directional information than ϕ_u, equations (2.5), (2.7) were approximated by P_N methods [9]. Sometimes, a few spherical harmonics will be enough to approximate ϕ_c well, the use of a low degree approximation allows then for the scattering operator to be approximated more efficiently. If a low degree P_N approximation is not enough, the accuracy of the method will be low. Another issue with the P_N approximation is the dense coupling of the boundary conditions, see Section 1.4. Unfortunately, the solution for collided and uncollided problems are generally inconsistent if different schemes, such as Monte-Carlo and P_N methods, are employed for ϕ_u and ϕ_c.

2.4 Relation of the source iteration method and the first collision source method

Before discussing the similarities between the source iteration method and the first collision source method we would like to remark that this comparison has been considered by other authors. For example in [13] an adaptive first collision source method is presented, and numerical examples for the source iteration method and FCSM are given and compared. It is shown that, while the methods are equivalent, the collision equations (2.5),(2.7) in the FCSM can be dicretized more efficiently with a quadrature of lower order than for the collision free equation.

It is easy to see that source iteration method and the first collision source method are similar techniques. To show this, let us consider the residual $R_k = \phi - \phi_k$. Subtracting equations (2.2) and (2.3) from (1.1) and (1.2) for $k = 1, 2$, we obtain

$$\begin{aligned} s \cdot \nabla_r R_2 + \sigma_t R_2 &= \sigma_s K R_1 \text{ in } R \times S^2, \\ R_2 &= 0 \text{ on } \Gamma_-. \end{aligned}$$

Combining the obtained system with (2.2) and (2.3) for $k = 1$, and taking $\phi_0 = 0$ as initial iterate we arrive at the system of the form (2.4)-(2.5),

$$\begin{aligned} s \cdot \nabla_r \phi_1 + \sigma_t \phi_1 &= q, \text{ on } R \times S^2, \\ s \cdot \nabla_r R_2 + \sigma_t R_2 &= \sigma_s K \phi_1 + \sigma_s K R_2, \text{ on } R \times S^2, \end{aligned}$$

with corresponding boundary conditions

$$\begin{aligned} \phi_1 &= g, \text{ on } \Gamma_-, \\ R_2 &= 0, \text{ on } \Gamma_-. \end{aligned}$$

Hence, we can draw the following conclusion for the source iteration method and the first collision source method: ϕ_1 corresponds to ϕ_u and R_2 corresponds to ϕ_c.

On the basis of the source iteration method and the first collision source method we can also construct another splitting method that we refer to as extended first collision source method.

2.5 Extended first collision source method

The observations regarding the first collision source method in Section 2.3 carry over if we use the following extension of the first collision source method, similar to [13]. We represent the function ϕ as a sum of M functions, where $M \geq 1$ (standard FCSM for $M = 1$):

$$\phi = \psi_0 + \psi_1 + ... + \psi_M. \tag{2.8}$$

In this formulation ψ_k, $k = 0, ..., M-1$ are the components of the solution that correspond to exactly k scattering events and ψ_M is a residual that corresponds to M or more scattering events.

Let L denote the following differential operator

$$L\psi(r,s) = s \cdot \nabla_r \psi(r,s) + \sigma_t(r)\psi(r,s)$$

and $K\psi(r,s) = \int_{S^2} k(s \cdot s')\psi(r,s')ds'$ be the integral scattering operator, see Section 1.2.

The components ψ_k, $0 \leq k \leq M-1$, then satisfy

$$\begin{aligned} &L\psi_0 = q \text{ on } R \times S^2, && \psi_0 = g \text{ on } \Gamma_-, \\ &L\psi_k = \sigma_s K\psi_{k-1} \text{ on } R \times S^2, && \psi_k = 0 \text{ on } \Gamma_-, \ k \geq 1. \end{aligned} \tag{2.9}$$

The closing equation for the residual is given by

$$L\psi_M = \sigma_s K\psi_M + \sigma_s K\psi_{M-1} \text{ on } R \times S^2, \qquad \psi_M = 0 \text{ on } \Gamma_-. \tag{2.10}$$

Observe that, if all the equations are added, we retrieve the original problem:

$$L\phi = q + \sigma_s K\phi \text{ on } R \times S^2, \qquad \phi = g \text{ on } \Gamma_-.$$

It is expected that the solution component ψ_k has less directional information than the previous one ψ_{k-1}. From a physical perspective the validity of the statement is quite obvious. After every scattering event with the medium the photons lose a fraction of the directional information. Indeed, suppose $\sigma_s = 1$ for simplicity, and let L_0 be the differential operator defined as

$$L_0\psi(r,s) = s \cdot \nabla_r \psi(r,s) + \sigma_t(r)\psi(r,s),$$

together with homogeneous boundary conditions. Using (2.9) we obtain for $k = 1, \ldots, M-1$ that

$$\psi_k = L_0^{-1} K\psi_{k-1}.$$

By iterating the argument, we therefore obtain for $k \geq 1$ that

$$\psi_k = L_0^{-1}(KL_0^{-1})^{k-1} K\psi_0.$$

Assume the eigendecomposition $KL_0^{-1}v_j = \lambda_j v_j$ with $|\lambda_j| < 1$. Since KL_0^{-1} is compact [11], we have that $\lambda_j \to 0$ as $j \to \infty$. For $u_j = L_0^{-1}v_j$, we observe that $\lambda_j L_0 u_j = Ku_j$. Expanding $K\psi_0$ as

$$K\psi_0 = \sum_{j=1}^{\infty} a_j v_j,$$

we then obtain that

$$\psi_k = L_0^{-1}(KL_0^{-1})^{k-1}K\psi_0 = L_0^{-1}\sum_{j=1}^{\infty} a_j\lambda_j^{k-1}v_j = \sum_{j=1}^{\infty} a_j\lambda_j^{k-1}u_j.$$

In view of Section 4.3, the functions u_j with λ_j close to 1 might be approximated well by Legendre polynomials of low degree, confirming the physical argument stated above.

Similar to the first collision source method the last equation (2.10) of the splitting system needs to be solved numerically, too, as it contains a term $K\psi_M(r,s)$. One possibility, as just argued, is to use the S_N or P_N method for the first $M-1$ equations and the P_N method (spherical harmonics approximation) of low degree for the closing M-th equation.

2.6 Conclusions

As a basic iterative technique in radiative transfer theory, the source iteration method (or equivalently the first collision source method) provides a good starting point for our research. Due to the consistency issues of the first collision source method we prefer the source iteration method. In Chapter 3 we will solve the isotropic radiative transfer equation and construct a robust preconditioner for the source iteration method. In Chapter 4 we will build upon the results obtained in Chapter 3 to construct highly efficient preconditioned iterative schemes, which are convergent in both the transport and diffusive regimes, also for anisotropic scattering problems.

Bibliography

[1] H. Egger, M. Schlottbom: An L^p theory for stationary radiative transfer. Applicable Analysis. An International Journal 93(6), 1283-1296 (2014)

[2] G. I. Marchuk, V. I. Lebedev: Numerical methods in the theory of neutron transport. Harwood Academic Publishers, Chur, London, Paris, New York (1986)

[3] W.H. Reed: New difference schemes for the neutron transport equation. Nuclear Science and Engineering, 46(2), 309-314 (1971)

[4] R.T. Ackroyd: A finite element method for neutron transport—I. Some theoretical considerations. Annals of Nuclear Energy, 5(2), 75-94 (1978)

[5] W.H. Reed: The effectiveness of acceleration techniques for iterative methods in transport theory. Nuclear Science and Engineering, 45(3), 245-254 (1971)

[6] M.L. Adams, E.W. Larsen: Fast iterative methods for discrete-ordinates particle transport calculations. Progress in nuclear energy, 40(1), 3-159 (2002)

[7] K. M. Case, P. F. Zweifel: Linear transport theory. Addison-Wesley, Reading (1967)

[8] J. C. Blake: Domain decomposition methods for nuclear reactor modelling with diffusion acceleration. Ph.D. thesis, University of Bath (2016)

[9] K.D. Lathrop: Remedies for ray effects. Nuclear Science and Engineering, 45(3), 255-268 (1971)

[10] R.E. Alcouffe: A first collision source method for coupling Monte Carlo and discrete ordinates for localized source problems. In Monte-Carlo Methods and Applications in Neutronics, Photonics and Statistical Physics, Springer, Berlin, Heidelberg, 352-366 (1985)

[11] F. Golse, P.L. Lions, B. Perthame, R. Sentis: Regularity of the moments of the solution of a transport equation. Journal of functional analysis, 76(1), 110-125 (1988)

[12] T.A. Wareing, J.E. Morel, D.K. Parsons: A first collision source method for ATTILA, an unstructured tetrahedral mesh discrete ordinates code. Los Alamos National Lab., NM, United States : N. p. (1998)

[13] W.J. Walters, A. Haghighat: The adaptive collision source method for discrete ordinates radiation transport. Annals of Nuclear Energy, 105, 45-58 (2017)

Chapter 3

On a convergent DSA preconditioned source iteration for a DGFEM method for radiative transfer

3.1 Introduction

We consider the numerical solution of the radiative transfer equation in plane parallel geometry, see Figure 1.1

$$\mu\partial_z\phi(z,\mu) + \sigma_t(z)\phi(z,\mu) = \frac{\sigma_s(z)}{2}\int_{-1}^{1}\phi(z,\mu')d\mu' + q(z,\mu), \qquad (3.1)$$

where $0 < z < Z$ and $-1 < \mu < 1$, and Z denotes the thickness of the slab and μ is the cosine of the polar angle of a unit vector. The function $\phi(z,\mu)$ models the equilibrium distribution of some quantity, like neutrons or photons [14, 13]. The basic physical principles embodied in (3.1) are transport, which is modeled by the differential operator $\mu\partial_z$, absorption with rate $\sigma_a = \sigma_t - \sigma_s$ and scattering with rate σ_s. Internal sources are described by the function q. In this work we will close the radiative transfer equation using inflow boundary conditions

$$\phi(0,\mu) = g^0(\mu) \quad \mu > 0, \quad \text{and} \quad \phi(Z,\mu) = g^Z(\mu) \quad \mu < 0. \qquad (3.2)$$

Such transport problems arise when the full three-dimensional model posed on $\mathbb{R}^2 \times (0,Z) \times \mathcal{S}^2$ obeys certain symmetries [13]. It has been studied in many instance due to simpler structure compared to three dimensional problems without symmetries; the methodology presented here directly carries over to the general case, see also the numerical examples presented below.

Classical deterministic discretization strategies are based on a semidiscretization in μ. One class of such strategies are the P_N-approximations, which are spectral methods based on truncated spherical harmonics expansions [21, 18], and we refer to [3, 24, 15] for variational discretization strategies using this approximation. The major advantage of P_N-approximations is that the scattering operator becomes diagonal. In addition, the matrix representation of the transport operator $\mu\partial_z$ is sparse. The main drawbacks of the P_N-method are that the variational incorporation of the inflow boundary condition

The content of this chapter was published in: O. Palii and M. Schlottbom, Computers & Mathematics with Applications, 79(12), pp. 3366-3377 (2020). https://doi.org/10.1016/j.camwa.2020.02.002

introduces a dense coupling of the spherical harmonics expansion coefficients making standard P_N-approximations quite expensive to solve. We note, however, that a modified variational formulation of the P_N-equations has recently been derived that leads to sparse matrices [6]. In any case, the success of spectral approximation techniques depends on the smoothness of the solution. In general, the solution ϕ is not smooth for $\mu = 0$, which is related to the inflow boundary conditions (3.2). Hence, the P_N-approximations will in general not converge spectrally. Resolving the non-smoothness of ϕ should improve the approximation considerably, and this observation led to the developments of double P_N-methods [18] or half space moment methods [23, 22], which are spectral methods on the intervals $\mu > 0$ and $\mu < 0$.

A second class of semidiscretizations are discrete ordinates methods that use a quadrature rule for the discretization of the μ-variable [13], with analysis provided in [16, 11]. Such methods are closely related to discontinuous Galerkin (DG) methods, see, e.g., [20, 2, 4, 5, 17, 26]. While allowing for local angular resolution, the main obstruction in the use of these methods is that the scattering operator leads to dense matrices, and a direct inversion of the resulting system is not possible in realistic applications. To overcome this issue, iterative solution techniques have been proposed. An often used iterative technique is Richardson iteration, i.e., the source iteration [19, 9, 36], but other Krylov space methods exist [1]. The key idea of the source iteration is to decouple scattering and transport, and to exploit that the transport part can be inverted efficiently. If $\sigma_s/\sigma_t \approx 1$, the convergence of these iterative methods is slow, and several preconditioning techniques have been proposed [5, 19, 9, 1]. Among the most used and simple preconditioners is the diffusion synthetic acceleration method (DSA), in which a diffusion problem is solved in every iteration. This is well motivated by asymptotic analysis [7, 8]. While Fourier analysis can be applied to special situations [19, 9], the convergence analysis is mainly open for the general case, i.e., for non-constant coefficients or non-periodic boundary conditions. A further complication in using the DSA preconditioner is that the resulting iterative scheme might diverge if the discretization of the diffusion problem and the discrete ordinates system are not consistent [9].

For isotropic scattering, integral methods for the approximation of the angular average of the solution have been proposed recently [25]. The efficiency of the iterative solver of [25] deteriorates if scattering becomes dominant, and a diffusion-based preconditioner is employed to reduce the number of iterations. Extensions to anisotropic scattering can be found in [27].

The contribution of this chapter is to develop discretizations that allow for local resolution of the non-smoothness of the solution, and which lead to discrete problems that can be solved efficiently by diffusion synthetic accelerated source iterations. Our approach builds upon an even-parity formulation of the radiative transfer equation derived from the mixed variational framework given in [15], where P_N-approximations have been treated in detail. We show that our DSA preconditioned source iteration converges already for the continuous problem. We present conforming *hp*-type approximation spaces for the discretization, and prove quasi-best approximation properties of the Galerkin approximation. In order to solve the resulting linear systems, we employ a DSA preconditioned Richardson iteration, which is just the infinite dimensional iteration projected to the approximation spaces. In particular, the finite di-

mensional iteration is guaranteed to converge for any discretization. Moreover, the inversion of the transport problem can be parallelized straightforwardly. In numerical experiments, even when employing low-order approximations, we observe that the developed method does not suffer from the ray effect, which is typically observed for discrete ordinates methods [18, 12].

The outline of the chapter is as follows: In Section 3.2 we introduce basic notation and recall the relevant function spaces. In Section 3.3 we introduce the even-parity equation for (3.1), show its well-posedness, and formulate an infinite dimensional DSA preconditioned source iteration, for which we show convergence. The approximation spaces are described in Section 3.4 and well-posedness of the Galerkin problems as well as quasi-best approximation results are presented. In Section 3.5 we discuss the efficient iterative solution of the resulting linear systems and provide a convergence proof for the discrete DSA scheme. Section 3.6 presents supporting numerical examples for a slab geometry and for multi-dimensional problems that show the good approximation properties of the proposed method as well as fast convergence of the iterative solver in multi-dimensional problems and in the diffusion limit. The chapter ends with some conclusions in Section 3.7.

3.2 Function spaces and further preliminaries

Following [10] we denote by $L^2(\mathcal{D})$ with $\mathcal{D} = (0, Z) \times (-1, 1)$ the usual Hilbert space of square integrable functions with inner product

$$(\phi, \psi) = \int_{-1}^{1} \int_{0}^{Z} \phi(z, \mu)\psi(z, \mu) dz d\mu$$

and induced norm $\|\phi\|_{L^2(\mathcal{D})} = (\phi, \phi)^{\frac{1}{2}}$. Furthermore, we define the Hilbert space

$$H_2^1(\mathcal{D}) = \{\phi \in L^2(\mathcal{D}) : \mu\partial_z\phi \in L^2(\mathcal{D})\}$$

of functions with square integrable weak derivatives with respect to the weighted differential operator $\mu\partial_z$ endowed with the corresponding graph norm.

In order to deal with boundary data, let us introduce the Hilbert space L^2_- that consists of measurable functions for which

$$\|\psi\|^2_{L^2_-} = \int_0^1 |\psi(0, \mu)|^2 |\mu| d\mu + \int_{-1}^0 |\psi(Z, \mu)|^2 |\mu| d\mu$$

is finite, and we denote by $\langle \psi, \phi \rangle_{L^2_-}$ the corresponding inner product on L^2_-. Similarly, L^2_+ denotes the space of outflow data. We have the following trace lemma [10].

Lemma 3.2.1. *If $\phi \in H_2^1(\mathcal{D})$, then there exist traces $\phi_{|\Gamma_-} \in L^2_-$ and $\phi_{|\Gamma_+} \in L^2_+$ and*

$$\|\phi_{|\Gamma_\pm}\|_{L^2_\pm} \leq \frac{C}{\sqrt{1 - e^{-Z}}} \|\phi\|_{H_2^1(\mathcal{D})}$$

with a constant $C > 0$ independent of ϕ and Z.

As a consequence of the trace lemma and the density of smooth functions in $H_2^1(\mathcal{D})$ the following integration-by-parts formula is true [10]

$$(\mu\partial_z\phi,\psi) = -(\phi,\mu\partial_z\psi) + \langle\phi,\psi\rangle_{L^2_+} - \langle\phi,\psi\rangle_{L^2_-}. \tag{3.3}$$

Throughout this chapter we make the following basic assumption:

(A1) $\sigma_s, \sigma_t \in L^\infty(0,Z)$ are non-negative and $\sigma_a = \sigma_t - \sigma_s \geq \gamma > 0$.

Assumption (A1) means that we consider absorbing media, which makes (3.1)-(3.2) well-posed [10].

Lemma 3.2.2. *Let assumption (A1) hold, and let $g \in L^2_-$ and $q \in L^2(\mathcal{D})$, then (3.1)-(3.2) has a unique solution $\phi \in H_2^1(\mathcal{D})$ that satisfies the a-priori bound*

$$\|\phi\|_{H_2^1(\mathcal{D})} \leq C(\|g\|_{L^2_-} + \|q\|_{L^2(\mathcal{D})}).$$

Assumption (A1) is not required to prove well-posedness for bounded geometries [30], or for slab problems with constant coefficients [28, Thm. 2.25].

Let $\mathcal{P} : L^2(\mathcal{D}) \to L^2(\mathcal{D})$ denote the L^2-projection onto constants in μ, i.e.,

$$(\mathcal{P}\psi)(z,\mu) = \frac{1}{2}\int_{-1}^{1}\psi(z,\mu')d\mu'.$$

Since $\sigma_t \in L^\infty(0,Z)$ is strictly positive, we can define the following norms on $L^2(\mathcal{D})$

$$\|\psi\|^2_{\sigma_t} = (\sigma_t\psi,\psi) \quad \text{and} \quad \|\psi\|^2_{\frac{1}{\sigma_t}} = (\frac{1}{\sigma_t}\psi,\psi). \tag{3.4}$$

Even-odd splitting

The even ϕ^+ and odd ϕ^- parts of a function $\phi \in L^2(\mathcal{D})$ are defined as

$$\phi^+(z,\mu) = \frac{1}{2}(\phi(z,\mu)+\phi(z,-\mu)), \qquad \phi^-(z,\mu) = \frac{1}{2}(\phi(z,\mu)-\phi(z,-\mu)).$$

Even-odd decompositions are frequently used in transport theory [21, 3]. Since, as functions of μ, even and odd functions are orthogonal in $L^2(-1,1)$, we can decompose $L^2(\mathcal{D})$ into orthogonal subspaces containing even and odd functions, respectively,

$$L^2(\mathcal{D}) = L^2(\mathcal{D})^+ \oplus L^2(\mathcal{D})^-.$$

Similarly, we will write $H_2^1(\mathcal{D})^\pm = H_2^1(\mathcal{D}) \cap L^2(\mathcal{D})^\pm$. As in [15], we observe that $\mu\partial_z\phi^\pm \in L^2(\mathcal{D})^\mp$ for any $\phi \in H_2^1(\mathcal{D})$, and $\mathcal{P}\phi^\pm \in L^2(\mathcal{D})^+$ for $\phi \in L^2(\mathcal{D})$. It turns out that the natural space for our formulation is

$$\mathbb{W} = H_2^1(\mathcal{D})^+ \oplus L^2(\mathcal{D})^-.$$

3.3 Weak formulation of the slab problem

Derivation

We follow the steps presented in [15] for multi-dimensional problems. The key idea is to rewrite the slab problem into a weak formulation for the even and odd parts of the solution. Multiplication of (3.1) with a test function $\psi \in \mathbb{W}^+$ and using orthogonality of even and odd functions gives that

$$(\mu\partial_z\phi^-, \psi^+) + ((\sigma_t - \sigma_s\mathcal{P})\phi^+, \psi^+) = (q^+, \psi^+).$$

Integration-by-parts (3.3) applied to the first term on the left-hand side yields that

$$(\mu\partial_z\phi^-, \psi^+) = -(\phi^-, \mu\partial_z\psi^+) + \langle\phi^-, \psi^+\rangle_{L^2_+} - \langle\phi^-, \psi^+\rangle_{L^2_-}.$$

Due to symmetries, we have that $\langle\phi^-, \psi^+\rangle_{L^2_+} = -\langle\phi^-, \psi^+\rangle_{L^2_-}$. Using (3.2), we have that $\phi^- = \phi - \phi^+ = g - \phi^+$ on the inflow boundary, which leads to

$$(\mu\partial_z\phi^-, \psi^+) = -(\phi^-, \mu\partial_z\psi^+) + 2\langle\phi^+, \psi^+\rangle_{L^2_-} - 2\langle g, \psi^+\rangle_{L^2_-}.$$

Thus, for any $\psi^+ \in \mathbb{W}^+$, it holds that

$$2\langle\phi^+, \psi^+\rangle_{L^2_-} - (\phi^-, \mu\partial_z\psi^+) + ((\sigma_t - \sigma_s\mathcal{P})\phi^+, \psi^+) = (q^+, \psi^+) + 2\langle g, \psi^+\rangle_{L^2_-}. \tag{3.5}$$

Testing (3.1) with an odd test function $\psi^- \in \mathbb{W}^-$, we obtain that

$$(\mu\partial_z\phi^+, \psi^-) + (\sigma_t\phi^-, \psi^-) = (q^-, \psi^-),$$

which implies that $\phi^- = \frac{1}{\sigma_t}(q^- - \mu\partial_z\phi^+) \in \mathbb{W}^-$. Using this expression for ϕ^- in (3.5), we deduce that $u = \phi^+$ is a solution to the following problem; cf. [15].

Problem 3.3.1. *Let $q \in L^2(\mathcal{D})$ and $g \in L^2_-$. Find $u \in \mathbb{W}^+$ such that*

$$a(u, v) = \ell(v) \quad \textit{for all } v \in \mathbb{W}^+. \tag{3.6}$$

where the bilinear form $a : \mathbb{W}^+ \times \mathbb{W}^+ \to \mathbb{R}$ is given by

$$a(u, v) = 2\langle u, v\rangle_{L^2_-} + (\frac{1}{\sigma_t}\mu\partial_z u, \mu\partial_z v) + ((\sigma_t - \sigma_s\mathcal{P})u, v), \tag{3.7}$$

and the linear form $\ell : \mathbb{W}^+ \to \mathbb{R}$ is defined as

$$\ell(v) = (q^+, v) + 2\langle g, \psi^+\rangle_{L^2_-} + (q^-, \frac{1}{\sigma_t}\mu\partial_z v). \tag{3.8}$$

Well-posedness

We endow $\mathbb{W}^+$ with the norm induced by the bilinear form a defined in (3.7), i.e.,

$$\|u\|_{\mathbb{W}} = \|u\|_a = a(u, u)^{\frac{1}{2}} \quad \text{for } u \in \mathbb{W}^+. \tag{3.9}$$

Using the Cauchy-Schwarz inequality, we obtain the following result.

Lemma 3.3.2. *The linear form $\ell : \mathbb{W} \to \mathbb{R}$ defined in* (3.8) *is bounded, i.e., for all $v \in \mathbb{W}^+$ it holds*

$$\ell(v) \leq (\|q^+\|^2_{\frac{1}{\sigma_a}} + \|q^-\|^2_{\frac{1}{\sigma_t}} + 2\|g\|^2_{L^2_-})^{\frac{1}{2}} \|v\|_a.$$

Since the space $\mathbb{W}^+$ endowed with the inner product induced by a is a Hilbert space, the unique solvability of Problem 3.3.1 is a direct consequence of the Riesz representation theorem.

Theorem 3.3.3. *Let Assumption (A1) hold true. Then Problem 3.3.1 has a unique solution $u \in \mathbb{W}^+$. Moreover, we have the bound*

$$\|u\|_a \leq (\|q^+\|^2_{\frac{1}{\sigma_a}} + \|q^-\|^2_{\frac{1}{\sigma_t}} + 2\|g\|^2_{L^2_-})^{\frac{1}{2}}.$$

Remark 3.3.4. *Setting $\phi^+ = u$ and $\phi^- = \frac{1}{\sigma_t}(q^- - \mu\partial_z u)$, one can show that $\mu\partial_z\phi^- \in L^2(\mathcal{D})$, i.e., $\phi = \phi^+ + \phi^- \in H^1_2(\mathcal{D})$ satisfies* (3.1) *in $L^2(\mathcal{D})$, cf. [15]. Using the trace lemma 3.2.1 and partial-integration* (3.3)*, we further can show that ϕ satisfies the boundary conditions* (3.2) *in the sense of traces. Hence, Theorem 3.3.3, independently, leads to a well-posedness result as Lemma 3.2.2 for* (3.1)-(3.2)*.*

DSA preconditioned source iteration in second order form

As a preparation for the numerical solution of the discrete systems that will be described below, let us discuss an iterative scheme in infinite dimensions for solving the radiative transfer equation. The basic idea is a standard one and consists of decoupling scattering and transport in order to compute successive approximations, viz., the source iteration [19, 9]. Next to the basic iteration, we describe a preconditioner which resembles diffusion synthetic acceleration (DSA) schemes using the notation of [9] or a KP_1 scheme using the terminology of [19].

In order to formulate the method, we introduce the following bilinear forms

$$k(u,v) = (\sigma_s \mathcal{P}u, v) \text{ and } b(u,v) = a(u,v) + k(u,v) \quad \text{for } u, v \in \mathbb{W}^+,$$

and denote the induced semi-norm and norm by $\|u\|_k$ and $\|u\|_b$, respectively.

The iteration scheme is defined as follows: For $u^k \in \mathbb{W}^+$ given, compute $u^{k+\frac{1}{2}} \in \mathbb{W}^+$ as the unique solution to

$$b(u^{k+\frac{1}{2}}, v) = k(u^k, v) + \ell(v) \quad \text{for all } v \in \mathbb{W}^+. \tag{3.10}$$

The half-step error $e^{k+\frac{1}{2}} = u - u^{k+\frac{1}{2}}$ satisfies

$$a(e^{k+\frac{1}{2}}, v) = k(u^{k+\frac{1}{2}} - u^k, v) \quad \text{for all } v \in \mathbb{W}^+. \tag{3.11}$$

The key idea is then to construct an easy-to-compute approximation to $e^{k+\frac{1}{2}}$ by Galerkin projection onto a suitable subspace $\mathbb{W}^+_1 \subset \mathbb{W}^+$. This approximation is then used to correct $u^{k+\frac{1}{2}}$ to obtain a more accurate approximation u^{k+1} to u. Define the following closed subspace of $\mathbb{W}^+$

$$\mathbb{W}^+_1 = \{u \in \mathbb{W}^+ : u = \mathcal{P}u\}, \tag{3.12}$$

i.e., $\mathbb{W}_1^+$ consists of functions in $\mathbb{W}^+$ that do not depend on μ. The correction $u_D^{k+\frac{1}{2}} \in \mathbb{W}_1^+$ is then computed by Galerkin projection of (3.11) to $\mathbb{W}_1^+$:

$$a(u_D^{k+\frac{1}{2}}, v) = k(u^{k+\frac{1}{2}} - u^k, v) \quad \text{for all } v \in \mathbb{W}_1^+, \tag{3.13}$$

and the new iterate is defined as

$$u^{k+1} = u^{k+\frac{1}{2}} + u_D^{k+\frac{1}{2}}. \tag{3.14}$$

If $u_D^{k+\frac{1}{2}}$ is a good approximation to $e^{k+\frac{1}{2}}$, then $e^{k+1} = e^{k+\frac{1}{2}} - u_D^{k+\frac{1}{2}}$ is small. The convergence proof for the iteration $\mathbb{W}^+ \to \mathbb{W}^+$, $u^k \mapsto u^{k+1}$ is based on spectral analysis [34, 35].

Lemma 3.3.5. *The half-step error $e^{k+\frac{1}{2}} = u - u^{k+1/2}$ of the iteration* (3.10) *satisfies*

$$\|e^{k+\frac{1}{2}}\|_a \leq c\|e^k\|_a,$$

with constant $c = \|\sigma_s/\sigma_t\|_\infty$.

Proof. We endow $\mathbb{W}^+$ with the inner product induced by the bilinear b in this proof, and define bounded, self-adjoint and positive operators A and K on $\mathbb{W}^+$ by

$$b(Au, v) = a(u, v), \quad b(Ku, v) = k(u, v) \quad \text{for } u, v \in \mathbb{W}^+.$$

Using $a = b - k$, we obtain that $A = I - K$. Using $u^{k+\frac{1}{2}} - u^k = e^k - e^{k+\frac{1}{2}}$, (3.11) can be written as

$$e^{k+\frac{1}{2}} = Ke^k.$$

We thus have that

$$\begin{aligned}\|e^{k+\frac{1}{2}}\|_a^2 = b((I-K)Ke^k, Ke^k) &= b(K^2(I-K)^{\frac{1}{2}}e^k, (I-K)^{\frac{1}{2}}e^k) \\ &\leq \max\sigma(K)^2\|(I-K)^{\frac{1}{2}}e^k\|_b^2 \leq c^2\|e^k\|_a^2.\end{aligned}$$

In the last step we have used the following bounds on the numerical range

$$0 \leq b(Kv, v) = k(v, v) \leq \|\frac{\sigma_s}{\sigma_t}\|_\infty(\sigma_t v, v) \leq cb(v, v),$$

which yields the spectral bounds $\sigma(K) \subset [0, c]$. □

Theorem 3.3.6. *Let Assumption (A1) hold, and let $c = \|\sigma_s/\sigma_t\|_\infty < 1$ be as in Lemma 3.3.5. For any $u^0 \in \mathbb{W}^+$, the iteration defined by* (3.10), (3.13), (3.14) *converges linearly to the solution u of Problem 3.3.1 with*

$$\|u - u^{k+1}\|_a \leq c\|u - u^k\|_a.$$

Proof. Since $u_D^{k+\frac{1}{2}}$ is the orthogonal projection of $e^{k+\frac{1}{2}}$ to $\mathbb{W}_1^+$ in the a-inner product it holds that

$$\|e^{k+1}\|_a = \|e^{k+\frac{1}{2}} - u_D^{k+\frac{1}{2}}\|_a = \inf_{v \in \mathbb{W}_1^+} \|e^{k+\frac{1}{2}} - v\|_a \leq \|e^{k+\frac{1}{2}}\|_a.$$

The assertion then follows from Lemma 3.3.5. □

Remark 3.3.7. *The convergence analysis presented in this section carries over verbatim to multi-dimensional problems without symmetries, and it can be extended immediately to more general (symmetric and positive) scattering operators. Moreover, if a Poincaré-Friedrichs inequality is available, cf., e.g., [24], then the case $\sigma_a \geq 0$ can be treated similarly as long as σ_t is uniformly bounded away from 0, and the source iteration converges also in this situation*

Remark 3.3.8. *Problem* (3.13) *is the weak formulation of the diffusion equation*

$$-\partial_z(\frac{1}{3\sigma_t}\partial_z u_D) + \sigma_a u_D = f \quad \text{in } (0, Z),$$

with $f = \sigma_s \mathcal{P}(u^{k+\frac{1}{2}} - u^k)$, complemented by Robin boundary conditions, which shows the close relationship to DSA schemes.

Remark 3.3.9. *The convergence analysis for the iteration without preconditioning, can alternatively be based on the following estimates. These estimates are the only ones in this paper that exploit that the scattering operator is related to the L^2-projector $\mathcal{P}$. First note that*

$$\|e^{k+\frac{1}{2}}\|_b^2 = \|\mathcal{P}e^{k+\frac{1}{2}}\|_{\sigma_t}^2 + \|(I-\mathcal{P})e^{k+\frac{1}{2}}\|_{\sigma_t}^2 + \|e^{k+\frac{1}{2}}\|_{L^2_-}^2 + \|\mu\partial_z e^{k+\frac{1}{2}}\|_{\frac{1}{\sigma_t}}^2$$

and that $\|e^{k+\frac{1}{2}}\|_b^2 = k(e^k, e^{k+\frac{1}{2}})$. Setting $c = \|\sigma_s/\sigma_t\|_\infty$, the Cauchy-Schwarz inequality yields

$$(1-\frac{\varepsilon}{2})\|\mathcal{P}e^{k+\frac{1}{2}}\|_{\sigma_t}^2 + \|(I-\mathcal{P})e^{k+\frac{1}{2}}\|_{\sigma_t}^2 + \|e^{k+\frac{1}{2}}\|_{L^2_-}^2 + \|\mu\partial_z e^{k+\frac{1}{2}}\|_{\frac{1}{\sigma_t}}^2 \leq \frac{c^2}{2\varepsilon}\|\mathcal{P}e^k\|_{\sigma_t}^2$$

for any $\varepsilon \in (0, 2]$. Choosing $\varepsilon = 2$ shows that parts of the error are smoothened independently of c, i.e.,

$$\|(I-\mathcal{P})e^{k+\frac{1}{2}}\|_{\sigma_t}^2 + \|e^{k+\frac{1}{2}}\|_{L^2_-}^2 + \|\mu\partial_z e^{k+\frac{1}{2}}\|_{\frac{1}{\sigma_t}}^2 \leq \frac{c^2}{4}\|\mathcal{P}e^k\|_{\sigma_t}^2,$$

while the angular average is hardly damped. In any case, this shows that $\mathcal{P}e^k$ converges to zero. It remains open how to exploit such an estimate to improve the analysis of the DSA preconditioned scheme above as it seems difficult to relate the latter smoothing property to the best approximation error of $e^{k+\frac{1}{2}}$ in the a-norm.

3.4 Galerkin approximations

In this section, we construct conforming approximation spaces $\mathbb{W}^+_{h,N} \subset \mathbb{W}^+$ in a two-step procedure. In a first step, we discretize the μ-variable using discontinuous ansatz functions. In a second step, we discretize the z-variable by continuous finite elements. Before stating the particular approximation space, we provide some general results. Let us begin with the definition of the discrete problem.

Problem 3.4.1. *Let $q \in L^2(\mathcal{D})$, $g \in L^2_-$, $\mathbb{W}^+_{h,N} \subset \mathbb{W}^+$ and let a and ℓ be defined as in Problem 3.3.1. Find $u_h \in \mathbb{W}^+_{h,N}$ such that for all $v \in \mathbb{W}^+_{h,N}$ there holds*

$$a(u_h, v_h) = \ell(v_h).$$

Since the bilinear form a induces the energy norm that we use in our analysis, we immediately obtain the following best approximation result.

Theorem 3.4.2. *Let $\mathbb{W}^+_{h,N}$ be a closed subspace of $\mathbb{W}^+$. Then, there exists a unique solution $u_h \in \mathbb{W}^+_{h,N}$ of Problem 3.4.1 that satisfies the a-priori estimate*

$$\|u_h\|_a \leq (\|q^+\|^2_{\frac{1}{\sigma_a}} + \|q^-\|^2_{\frac{1}{\sigma_t}} + 2\|g\|^2_{L^2_-})^{\frac{1}{2}}, \tag{3.15}$$

and the following best approximation error estimate

$$\|u - u_h\|_a = \inf_{v_h \in \mathbb{W}^+_{h,N}} \|u - v_h\|_a. \tag{3.16}$$

In the next sections, we will discuss some particular discretizations. We note that these generalize the spherical harmonics approach presented in [15].

hp-type semidiscretization in μ

Since we consider even functions, we require that the partition of the interval $[-1, 1]$ for the μ variable respects the point symmetry $\mu \mapsto -\mu$. For simplicity, we thus partition the interval $[0, 1]$ only, and the partition of $[-1, 0]$ is implicitly defined by reflection.

Let $N \in \mathbb{N}$ be a positive integer, and define intervals $\bar{\mu}_n = (\mu_{n-\frac{1}{2}}, \mu_{n+\frac{1}{2}})$, $n = 1, \ldots, N$, such that $\mu_{\frac{1}{2}} = 0$ and $\mu_{N+\frac{1}{2}} = 1$, and set $\Delta\mu_n = \mu_{n+\frac{1}{2}} - \mu_{n-\frac{1}{2}}$ and $\mu_n = (\mu_{n+\frac{1}{2}} + \mu_{n-\frac{1}{2}})/2$. Denote $\chi_n(\mu)$ the characteristic function of the interval $\bar{\mu}_n$. For $\mu > 0$, we define the piecewise functions

$$Q_{n,l}(\mu) = \sqrt{\frac{2l+1}{2}} P_l\Big(2\frac{\mu - \mu_{n-\frac{1}{2}}}{\Delta\mu_n} - 1\Big)\chi_n(\mu), \quad \mu > 0,$$

where P_l denotes the lth Legendre polynomial. Hence, $\{Q_{n,l}\}_{l=0}^L$ is an L^2-orthogonal basis for the space of polynomials of degree L on each interval $\bar{\mu}_n$. For $\mu > 0$, we set $Q^{\pm}_{n,l}(\mu) = Q_{n,l}(\mu)$, and for $\mu < 0$, we set $Q^{\pm}_{n,l}(\mu) = \pm Q^{\pm}_{n,l}(-\mu)$. The semidiscretization of the even parts is then

$$u(z, \mu) \approx u_h(z, \mu) = \sum_{n=1}^{N} \sum_{l=0}^{L} \phi^+_{n,l}(z) Q^+_{n,l}(\mu).$$

Remark 3.4.3. *If we partition the interval $[-1, 1]$ for the angular variable by a single element, we obtain truncated spherical harmonics approximations, see, e.g., [18, 15]. Partitioning of $[-1, 1]$ into two symmetric intervals $(-1, 0) \cup (0, 1)$ corresponds to the double P_L-method [18], which generalizes in multiple dimensions to half space moment methods [23]. The latter can resolve the non-smoothness of ϕ at $\mu = 0$, and, thus might yield spectral convergence on the intervals $\mu > 0$ and $\mu < 0$.*

Fully discrete scheme

In order to obtain a conforming discretization, we approximate the coefficient functions $\phi^+_{n,l}$ using $H^1(0,Z)$-conforming elements. Let $J \in \mathbb{N}$, and $\bar{z}_j = (z_{j-1}, z_j)$ be such that

$$[0,Z] = \cup_{j=1}^{J} \mathrm{clos}(\bar{z}_j)$$

is a partition of $(0,Z)$. Let $p \geq 1$ and denote $\mathbb{P}_p$ the space of polynomials of degree p. The full approximation space is then defined by

$$\mathbb{W}^+_{h,N} = \{\psi^+_h(z,\mu) = \sum_{n=1}^{N}\sum_{l=0}^{L} \psi^+_{n,l}(z) Q^+_{n,l}(\mu) : \psi^+_{n,l}(z) \in H^1(0,Z), \quad \psi^+_{n,l|\bar{z}_j} \in \mathbb{P}_p\}. \tag{3.17}$$

The choice (3.17) for the approximation space $\mathbb{W}^+_{h,N}$ corresponds to a *hp*-type finite element method, for which the assertion of Theorem 3.4.2 holds true.

Remark 3.4.4. *If the solution $u_h \in \mathbb{W}^+_{h,N}$ to Problem 3.4.1 is computed, we compute even and odd approximations to the solution ϕ of (3.1) via $\phi^+_h = u_h$ and the odd part $\phi^-_h \in \mathbb{W}^-_h$ as the solution to the variational problem $(\phi^-_h, \psi^-_h) = (\frac{1}{\sigma_t}(q^- - \mu\partial_z u_h), \psi^-_h)$ for all $\psi^-_h \in \mathbb{W}^-_h$, where*

$$\mathbb{W}^-_{h,N} = \{\psi^-_h(z,\mu) = \sum_{n=1}^{N}\sum_{l=0}^{L+1} \psi^-_{n,l}(z) Q^-_{n,l}(\mu) : \psi^-_{n,l|\bar{z}_j} \in \mathbb{P}_{p-1}\}.$$

We note that this space satisfies the compatibility condition $\mu\partial_z \mathbb{W}^+_{h,N} \subset \mathbb{W}^-_{h,N}$, which makes this pair of approximation spaces suitable for a direct approximation of a corresponding mixed formulation, cf. [15]. The even-parity formulation that we consider here corresponds then to the Schur complement of the mixed problem, cf. [15]. The reader should note the different degrees in the polynomial approximations, e.g., if ϕ^+_h is piecewise constant in angle, then ϕ^-_h is piecewise linear.

3.5 Discrete preconditioned source iteration

In order to solve the discrete variational problem defined in Problem 3.4.1, we proceed as in Section 3.3, but with $\mathbb{W}^+$ and $\mathbb{W}^+_1$ replaced by $\mathbb{W}^+_{h,N}$ and $\mathbb{W}^+_{h,1}$, respectively. We note that $\mathbb{W}^+_{h,1} \subset \mathbb{W}^+_1$ consists of functions in $\mathbb{W}^+_{h,N}$ that do not dependent on μ.

The finite dimensional counterpart of the DSA preconditioned source iteration is then defined as follows: For given $u^k_h \in \mathbb{W}^+_{h,N}$, compute $u^{k+\frac{1}{2}}_h \in \mathbb{W}^+_{h,N}$ as the unique solution to

$$b(u^{k+\frac{1}{2}}_h, v_h) = k(u^k_h, v_h) + \ell(v_h) \quad \text{for all } v_h \in \mathbb{W}^+_{h,N}. \tag{3.18}$$

The correction $u^{k+\frac{1}{2}}_{h,D} \in \mathbb{W}^+_{h,1}$ is defined by Galerkin projection of $e^{k+\frac{1}{2}}_h$ to $\mathbb{W}^+_{h,1}$:

$$a(u^{k+\frac{1}{2}}_{h,D}, v_h) = k(u^{k+\frac{1}{2}}_h - u^k_h, v_h) \quad \text{for all } v_h \in \mathbb{W}^+_{h,1}, \tag{3.19}$$

and the new iterate is defined as

$$u_h^{k+1} = u_h^{k+\frac{1}{2}} + u_{h,D}^{k+\frac{1}{2}}. \tag{3.20}$$

Using the same arguments as above, we obtain the following convergence result.

Theorem 3.5.1. *Let Assumption (A1) hold, and let $c < 1$ be as in Lemma 3.3.5. For any $u_h^0 \in \mathbb{W}_{h,N}^+$, the iteration defined by* (3.18), (3.19), (3.20) *converges linearly with*

$$\|u_h - u_h^{k+1}\|_a \leq c\|u_h - u_h^k\|_a.$$

Remark 3.5.2. *Similar to Remark 3.3.8, $u_{h,D}^{k+\frac{1}{2}}$ is the Galerkin projection to $\mathbb{W}_{h,1}^+$ of the weak solution to*

$$-\partial_z(D(z)\partial_z u_D) + \sigma_a u_D = f, \quad 0 < z < Z,$$

with μ-grid dependent diffusion coefficient $D(z)$ and $f = \sigma_s \mathcal{P}(u_h^{k+\frac{1}{2}} - u_h^k)$. If piecewise constant functions in angle are employed, then $D(z) = \frac{1}{3\sigma_t(z)}(1 + \frac{1}{4}\sum_{n=1}^N \Delta\mu_n^3)$.

Remark 3.5.3. *Once the scattering term in the right-hand side of* (3.18) *has been computed, the half-step iterate $u_h^{k+\frac{1}{2}}$ can be computed independently for each direction, and, thus, its computation can be parallelized.*

3.6 Numerical examples

In this section, we report on the accuracy of the proposed discretization scheme and its efficient numerical solution using the DSA preconditioned source iteration of Section 3.5. We restrict our discussion to a low-order method that offers local resolution. To do so, we fix $p = 1$ and $L = 0$ in the definition of $\mathbb{W}_{h,N}^+$, while N might be large. Hence, the approximation space $\mathbb{W}_{h,N}^+$ consists of discontinuous, piecewise constant functions in the angular variable and continuous, piecewise linear functions in z.

Manufactured solutions

To investigate the convergence behavior, we use manufactured solutions, i.e., the exact solution is

$$\phi(z,\mu) = |\mu| e^{-\mu} e^{-z(1-z)}, \tag{3.21}$$

with parameters $\sigma_a = 1/100$, $\sigma_s(z) = 2 + \sin(\pi z)/2$ and $Z = 1$, and source terms defined accordingly. We computed the numerical solution u_h using the DSA preconditioned iteration (3.18), (3.19), (3.20). We stopped the iteration using the a-posteriori stopping rule

$$\|u_h^k - u_h^{k-1}\|_a \leq \varepsilon,$$

where $\varepsilon = 10^{-10}$ is a chosen tolerance. The approximations of the even and odd parts of the solutions are recovered as described in Remark 3.4.4.

For the chosen parameters, the predicted convergence rate is $c \approx 0.996$. Table 3.1 shows the errors for different discretization parameters N and J. As expected we observe a linear rate of convergence with respect to N, cf. Theorem 3.4.2. In addition, we observe that the preconditioned source iteration converged within at most 15 iterations, and the expression $\|u_h^k - u_h^{k-1}\|_a$ decreased by 0.21 in each iteration. The convergence rate with respect to J is initially quadratic, which is better than predicted by Theorem 3.4.2, and then saturates for fixed $N = 8192$. As before, the preconditioned source iteration converged within at most 15 iterations with a decrease by a factor of 0.21 of $\|u_h^k - u_h^{k-1}\|_a$.

The observed convergence rate is thus as the one obtained by classical Fourier analysis for constant coefficients and periodic boundary conditions, which is bounded by 0.2247 [9]. Furthermore, we observe that the rate of convergence does not depend on the grid size as predicted by Theorem 3.5.1, cf. Remark 3.5.2, i.e., using the terminology of [9], our discrete diffusion approximation is consistently discretized.

N	E_h	rate
512	1.61e-04	
1024	8.07e-05	0.99
2048	4.04e-05	0.99
4096	2.04e-05	0.99
8192	1.06e-05	0.95

J	E_h	rate
16	7.88e-04	
32	1.99e-04	1.98
64	5.14e-05	1.96
128	1.63e-05	1.66
256	1.06e-05	0.62

Table 3.1: Observed errors $E_h = \|\phi - \phi_h\|_{L^2(\mathcal{D})}$ between finite element solution ϕ_h and the manufactured solution ϕ defined in (3.21) together with the rate of convergence of E_h. Left: Convergence for different discretization parameters N, and $J = 256$. Right: Convergence for different discretization parameters J and $N = 8192$.

In the next section, we present a more detailed study of spectrum of the preconditioned operator.

Eigenvalue studies of the preconditioned operator

In this section, we consider the spectrum of the error propagation operator $\mathcal{P}e_h^n \mapsto \mathcal{P}e_h^{n+1}$, where $\mathcal{P}$ is the L^2-projection onto constants in angle defined in Section 3.2 and e_h^n is the sequence of errors generated by the DSA preconditioned source iteration, cf. Remark 3.3.9. The function $\mathcal{P}e_h^n$ depends only on z, and, assuming $\sigma_s > 0$, we can measure the projected error using the norm induced by σ_s, i.e.,

$$\|\mathcal{P}e_h^n\|_{\sigma_s} = \|e_h^n\|_k.$$

We choose the following scattering and absorption parameters

$$\sigma_s(z) = \begin{cases} 2 + \sin(2\pi z), & z \leq \frac{1}{2} \\ 102 + \sin(2\pi z), & z > \frac{1}{2}, \end{cases} \qquad \sigma_a(z) = \begin{cases} 10.01, & z \leq \frac{1}{2} \\ 0.01, & z > \frac{1}{2}. \end{cases}$$

We notice that both parameters have huge jumps and that the predicted convergence rate is $c = \|\sigma_s/\sigma_t\|_\infty \approx 0.9999$, cf. Theorem 3.3.6.

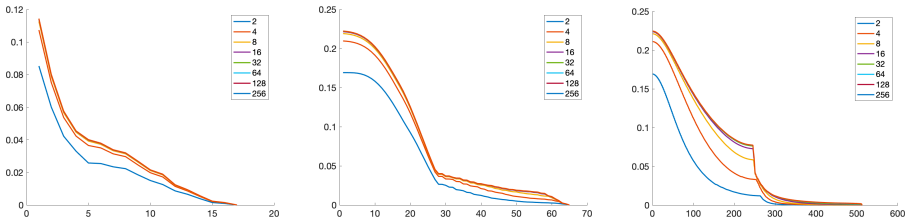


Figure 3.1: Spectra of the error propagation operator $\mathcal{P}e_h^n \mapsto \mathcal{P}e_h^{n+1}$ for different spatial discretizations $J = 16, 64, 512$ (from left to right). Each plot contains the corresponding spectra for $N = 2^i$, $i = 1, \dots, 8$.

In Figure 3.1 we plot the spectrum of the error propagation operator for different mesh sizes. All eigenvalues are bounded from above by 0.2247, which is inline with the results of [19, 9]. For each spatial discretization, we observe that the corresponding eigenvalues are monotonically increasing with N. The spectra for $N \geq 16$ lie on top of each other, indicating convergence of the eigenvalues. In particular, also for the coarse grid approximations the spectrum is uniformly bounded by 0.2247, again confirming the robustness of the method with respect to different approximations.

Multi-dimensional problems

Although the theory has been presented for slab problems, it becomes clear from our proofs that the results carry over verbatim to truly multi-dimensional problems, for which the radiative transfer equation writes as

$$s \cdot \nabla_x \phi(x, s) + \sigma_t(x)\phi(x, s) = \sigma_s(x)\mathcal{P}\phi + q(x, s).$$

In the following, let us consider homogeneous inflow boundary conditions, and homogeneity of the problem with respect to one spatial variable, i.e., we assume that the solution depends only on $x \in \mathcal{R} \subset \mathbb{R}^2$ and $s \in \mathbb{S}^2 \subset \mathbb{R}^3$.

We discretize $L^2(\mathbb{S}^2)$ by approximating $\mathbb{S}^2$ using a geodesic polyhedron consisting of flat triangles, see Figure 3.2. The approximation space in angle then consists of standard discontinuous finite element spaces associated to this triangulation. In the following, we focus on piecewise constant approximations, but higher order approximations are straightforward if the geometry approximation is also of higher order. The next paragraph is concerned with the accuracy of our method.

Manufactured solutions Let $\mathcal{R} = (0, 1) \times (0, 1)$ be the unit square, and let $\sigma_s = 2$ and $\sigma_a = 0.01$. Define the source function q such that

$$\phi(x, s) = \sin(\pi x_1) \sin(\pi x_2)(1 + s_1 + s_2^2 + s_3^3) \tag{3.22}$$

is the exact solution. In Table 3.2 we report on the numerical errors for different computational grids.

As expected from our error estimates, we observe linear convergence of the error in terms of the mesh size until saturation occurs. Note that for $N = 4096$ and $J = 4225$ the number of degrees of freedom is 17 305 600. The proven

N	E_h	rate
4	6.55e-01	
16	3.48e-01	0.91
64	1.87e-01	0.90
256	1.06e-01	0.81
1024	7.37e-02	0.53

J	E_h	rate
25	6.89e-01	
81	4.12e-01	0.75
289	2.25e-01	0.87
1089	1.18e-01	0.93
4225	6.30e-02	0.91

Table 3.2: Observed errors $E_h = \|\phi^+ - \phi_h^+\|_a$ between finite element solution ϕ_h and the manufactured solution ϕ defined in (3.22) together with the rate of convergence of E_h. Left: Convergence for different discretization parameters N, and $J = 4225$ vertices. Right: Convergence for different discretization parameters J and $N = 4096$ triangles on a half-sphere.

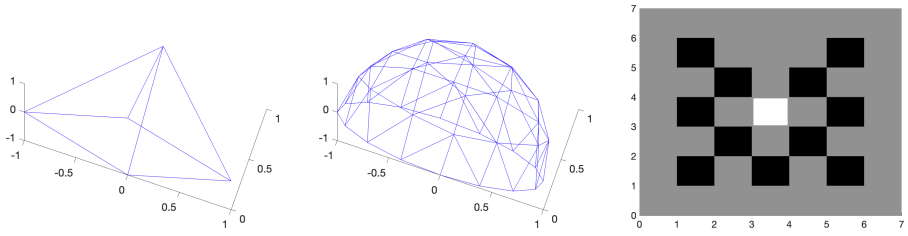


Figure 3.2: Left and middle: Approximation of the half-sphere with $N = 4$ and $N = 64$ triangles. Right: Geometry of the lattice problem.

contraction rate for the iteration is $c = 0.995$, while the observed maximal value of $\|u_h^k - u_h^{k-1}\|_a / \|u_h^{k-1} - u_h^{k-2}\|_a$, i.e., the minimal reduction rate, was 0.2 indicating much faster convergence of the iteration. After at most 16 iterations the stopping criterion $\|u_h^k - u_h^{k-1}\|_a \leq \varepsilon$ was reached.

The lattice problem A non-smooth test case without analytical solution is the lattice problem [12], which models the core of a neutron reactor. The computational domain is a square $\mathcal{R} = (0,7) \times (0,7)$. The absorption and scattering rates are piecewise constant functions. We define $\sigma_a = 10$ in the black regions shown in Figure 3.2, and $\sigma_a = 0$ elsewhere. We set $\sigma_s = 1$ in the grey and white regions and $\sigma_s = 0$ elsewhere. The source is defined by $q(x,s) = 1$ in the white region and $q(x,s) = 0$ elsewhere. Note that due to the availability of a Poincaré-Friedrichs inequality [24], the case $\sigma_a = 0$ leads to a well-posed radiative transfer problem, and, since $\sigma_s + \sigma_a \geq 1$, the theory presented here is applicable. Moreover, the constant c will depend on the constant from the Poincaré-Friedrichs inequality and $c < 1$.

In Figure 3.3 we show the angular averages of the computed solutions for two different grids with $J = 9\,801$ vertices in the spatial grid and $N = 4$ triangles on a half-sphere and for $J = 78\,961$ and $N = 64$, respectively. We note that our solutions do not suffer from ray effects, cf. [18, 12]. The preconditioned source iteration converged with 9 and 17 iterations for the coarse grid approximation and for the fine grid approximation, respectively. This amounts to an error reduction per iteration of at least 0.04 for the coarse grid discretization and of at least 0.2 for the fine grid discretization.

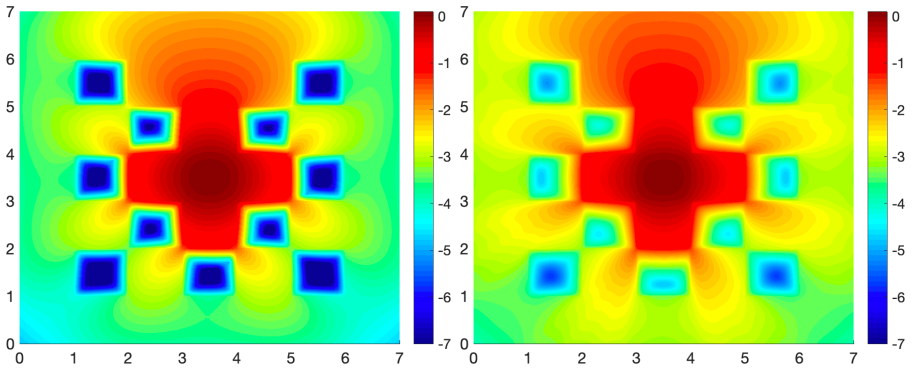


Figure 3.3: Angular average of the computed solution in a $\log_{10}$-scale for the lattice problem for $J = 9\,801$ spatial vertices and $N = 4$ triangles on a half-sphere (left) and $J = 78\,961$ spatial vertices and $N = 64$ triangles on a half-sphere (right).

Diffusion scaling Next, let us investigate the behavior of the preconditioned source iteration for scaled parameters. Introducing a scale parameter $\delta > 0$, a diffusion limit is obtained by the scaling [29]

$$\frac{\bar{\sigma}_s(x)}{\delta}, \quad \delta\bar{\sigma}_a(x), \quad \delta\bar{q}(x),$$

when both parameters $\bar{\sigma}_s$ and $\bar{\sigma}_a$ are bounded and strictly bounded away from zero. Since this is not the case for the lattice problem, we consider the following parameters

$$\sigma_s^\delta(x) = \frac{\sigma_s(x) + 1/10}{\delta}, \qquad \sigma_a^\delta = \delta(\sigma_a(x) + 1/10), \qquad q^\delta(x,s) = \delta q(x).$$

For $\delta \to 0$, the corresponding solution u^δ will converge to the solution of a diffusion problem; for non-smooth coefficients see [8]. The parameter c defined in Lemma 3.3.5 is bounded by $O(1/\delta)$. Table 3.3 shows the iteration counts

	$J = 9\,801$				$J = 78\,961$			
	$N = 4$		$N = 64$		$N = 4$		$N = 64$	
$1/\delta$	k	rate	k	rate	k	rate	k	rate
1	9	0.04	15	0.16	9	0.04	15	0.17
10	9	0.06	15	0.22	9	0.06	16	0.25
100	8	0.06	13	0.22	9	0.07	15	0.27
1000	5	0.01	7	0.06	6	0.05	10	0.17

Table 3.3: Iteration counts k and minimal reduction rates for $\|u_h^k - u_h^{k-1}\|_a$ for the lattice problem with scaled parameters σ_s^δ, σ_a^δ and q^δ for different δ and discretizations with N triangles on a half-sphere and J vertices in the spatial mesh.

and the minimum reduction of $\|u_h^k - u_h^{k-1}\|_a$ during the iteration. We observe that the preconditioned iteration is robust with respect to $\delta \to 0$ for different

meshes. The convergence rate on the finest grid is, however, slightly worse than the convergence rate for slab problems.

3.7 Conclusions

We investigated discontinuous angular and continuous spatial approximations of the even-parity formulation for the radiative transfer equation. Certain instances of these approximations are closely related to classical discretizations, such as truncated spherical harmonics approximations, double P_N-methods or discrete ordinates methods.

We considered a diffusion accelerated preconditioned source iteration for the solution of the resulting variational problems that has been formulated in infinite dimensions. Convergence rates of this iteration have been proven. The Galerkin approach used for the discretization allowed to translate the results for the infinite dimensional iteration directly to the discrete problems. In particular, the discrete iteration converges independently of the chosen resolution.

The theoretically proven convergence rate in Theorem 3.3.6 is not robust in the limit of large scattering, while numerical results show that in practice the preconditioned iteration converges robustly even in scattering dominated problems. One approach to obtain an improved convergence rates estimate is to estimate the best approximation error $\inf_{v\in\mathbb{W}_1^+} \|e^{k+\frac{1}{2}} - v\|_a$, which, however, seems rather difficult, and we postpone a corresponding rigorous analysis to future research.

The term diffusion acceleration is linked to the usage of the space $\mathbb{W}_1^+$ in (3.12) that consists of functions constant in angle. This choice is, however, not essential, and other closed subspaces can be employed, which allows for the construction of multi-level schemes, cf. also [5, 31, 32, 33]. The multi-level approach might lead to a feasible approach to estimate the best approximation error.

In any case, the DSA preconditioner can be combined with a conjugate gradient method in order to further reduce the number of iterations. Moreover, unlike many discrete ordinates schemes, the numerical approximation did not suffer from the ray effect in our numerical examples. We postpone a detailed study of this phenomenon to future research.

Bibliography

[1] J. S. Warsa, T. A. Wareing, J. E. Morel: Krylov iterative methods and the degraded effectiveness of diffusion synthetic acceleration for multidimensional S_N calculations in problems with material discontinuities. Nuclear Science and Engineering 147 (3) (2004) 218–248 (2004). doi:10.13182/NSE02-14

[2] T. Wareing, J. McGhee, J. Morel, S. Pautz: Discontinuous finite element S_N methods on three-dimensional unstructured grids, Nuclear Science and Engineering 138 (2001) 1–13 (2001)

[3] J. Pitkäranta: Approximate solution of the transport equation by methods of Galerkin type. J. Math. Anal. and Appl. 60, 186–210 (1977)

[4] J. C. Ragusa, J.-L. Guermond, G. Kanschat: A robust S_N-DG-approximation for radiation transport in optically thick and diffusive regimes. J. Comput. Phys. 231 (4), 1947–1962 (2012). doi:10.1016/j.jcp.2011.11.017

[5] G. Kanschat, J.-C. Ragusa: A robust multigrid preconditioner for S_NDG approximation of monochromatic, isotropic radiation transport problems. SIAM J. Sci. Comput. 36 (5), A2326–A2345 (2014). doi:10.1137/13091600X

[6] H. Egger, M. Schlottbom: A perfectly matched layer approach for radiative transfer in highly scattering regimes., SIAM J. Numer. Anal. 57 (5), 2166–2188 (2019)

[7] E. W. Larsen, J. B. Keller: Asymptotic solution of neutron transport problems for small mean free paths. J. Mathematical Phys. 15, 75–81 (1974). doi:10.1063/1.1666510

[8] H. Egger, M. Schlottbom: Diffusion asymptotics for linear transport with low regularity, Asymptot. Anal. 89 (3-4), 365–377 (2014)

[9] M. L. Adams, E. W. Larsen: Fast iterative methods for discrete-ordinates particle transport calculations. Progress in Nuclear Energy 40 (1), 3–159 (2002)

[10] V. Agoshkov: Boundary value problems for transport equations. Modeling and Simulation in Science, Engineering and Technology, Birkhäuser, Boston (1998)

[11] M. Asadzadeh: Analysis of a fully discrete scheme for neutron transport in two-dimensional geometry. SIAM J. Numer. Anal. 23, 543–561 (1986)

[12] T. A. Brunner: Forms of approximate radiation transport. In: Nuclear Mathematical and Computational Sciences: A Century in Review, A Century Anew Gatlinburg, American Nuclear Society, LaGrange Park, IL, 2003, Tennessee, April 6-11 (2003)

[13] K. M. Case, P. F. Zweifel: Linear transport theory. Addison-Wesley, Reading (1967)

[14] S. Chandrasekhar: Radiative Transfer. Dover Publications, Inc. (1960)

[15] H. Egger, M. Schlottbom: A mixed variational framework for the radiative transfer equation. Math. Mod. Meth. Appl. Sci. 22, 1150014 (2012)

[16] C. Johnson, J. Pitkäranta: Convergence of a fully discrete scheme for two-dimensional neutron transport. SIAM J. Numer. Anal. 20, 951–966 (1983)

[17] J. Kópházi, D. Lathouwers: A space-angle DGFEM approach for the Boltzmann radiation transport equation with local angular refinement. J.Comput. Phys. 297, 637–668 (2015). doi:10.1016/j.jcp.2015.05.031

[18] E. E. Lewis, W. F. Miller Jr.: Computational methods of neutron transport. John Wiley & Sons, Inc., New York Chichester Brisbane Toronto Singapore (1984)

[19] G. I. Marchuk, V. I. Lebedev: Numerical methods in the theory of neutron transport. Harwood Academic Publishers, Chur, London, Paris, New York (1986)

[20] W. H. Reed, T. R. Hill: Triangular mesh methods for the neutron transport equation. Tech. rep., Los Alamos Scientific Laboratory of the University of California (1973)

[21] V. S. Vladimirov: Mathematical problems in the one-velocity theory of particle transport. Tech. rep., Atomic Energy of Canada Ltd. AECL-1661. translated from Transactions of the V.A. Steklov Mathematical Institute (61) (1961)

[22] M. Frank: Approximatemodels for radiative transfer. Bull. Inst. Math. Acad. Sin. (N.S.) 2 (2), 409–432 (2007)

[23] B. Dubroca, M. Frank, A. Klar, G. Thömmes: A half space moment approximation to the radiative heat transfer equations. ZAMM Z. Angew. Math. Mech. 83 (12), 853–858 (2003). doi:10.1002/zamm.200310055

[24] T. A. Manteuffel, K. J. Ressel, G. Starke: A boundary functional for the least-squares finite-element solution for neutron transport problems. SIAM J.Numer. Anal. 2, 556–586 (2000)

[25] K. Ren, R. Zhang, Y. Zhong: A fast algorithm for radiative transport in isotropic media. Journal of Computational Physics 399 (2019) 108958 (2019). doi:https://doi.org/10.1016/j.jcp.2019.108958.

[26] P. Clarke, H. Wang, J. Garrard, R. Abedi, S. Mudaliar: Space-angle discontinuous Galerkin method for plane-parallel radiative transfer equation. Journal of Quantitative Spectroscopy and Radiative Transfer 233, 87 – 98 (2019). doi:https://doi.org/10.1016/j.jqsrt.2019.02.027

[27] Y. Fan, J. An, L. Ying: Fast algorithms for integral formulations of steady-state radiative transfer equation. Journal of Computational Physics 380, 191–211 (Mar 2019). doi:10.1016/j.jcp.2018.12.014

[28] J. C. Blake: Domain decomposition methods for nuclear reactor modelling with diffusion acceleration. Ph.D. thesis, University of Bath (2016)

[29] R. Dautray, J. L. Lions: Mathematical analysis and numerical methods for science and technology. Evolution Problems II, Vol. 6, Springer, Berlin (1993)

[30] H. Egger, M. Schlottbom: An L^p theory for stationary radiative transfer. Appl. Anal. 93 (6), 1283–1296 (2014). doi:10.1080/00036811.2013.826798

[31] J. E. Morel, T. A. Manteuffel: An angular multigrid acceleration technique for S_N equations with highly forward-peaked scattering. Nuclear Science and Engineering 107 (4), 330–342 (Apr 1991). doi:10.13182/nse91-a23795

[32] B. Lee: A multigrid framework for S_N discretizations of the Boltzmann transport equation. SIAM Journal on Scientific Computing 34 (4), A2018–A2047 (Jan 2012). doi:10.1137/110841199

[33] J. D. Densmore, D. F. Gill, J. M. Pounders: Cellwise block iteration as a multigrid smoother for discrete-ordinates radiation-transport calculations. Journal of Computational and Theoretical Transport 46 (1), 20–45 (Jun 2016). doi:10.1080/23324309.2016.1161650

[34] G. Helmberg: Introduction to spectral theory in Hilbert space. North-Holland Series in Applied Mathematics and Mechanics, Vol. 6, North-Holland Publishing Co., Amsterdam-London; Wiley Interscience Division John Wiley & Sons, Inc., New York (1969)

[35] D. Werner: Funktionalanalysis. extended Edition, Springer-Verlag, Berlin (2000).

[36] W. Dahmen, F. Gruber, O. Mula: An adaptive nested source term iteration for radiative transfer equations. Mathematics of Computation 1 (Nov 2019). doi:10.1090/mcom/3505

Chapter 4

On robustly convergent and efficient iterative methods for anisotropic radiative transfer

4.1 Introduction

Radiative transfer models describe the streaming, absorption, and scattering of radiation waves propagating through a turbid medium occupying a bounded convex domain $R \subset \mathbb{R}^d$, and they arise in a variety of applications, e.g., neutron transport [1, 2], heat transfer [43], climate sciences [48], geosciences [27] or medical imaging and treatment [28, 30, 29]. The underlying physical model can be described by the anisotropic radiative transfer equation,

$$s \cdot \nabla_r u(s,r) + \sigma_t(r)u(s,r) = \sigma_s(r) \int_S k(s \cdot s')u(s',r)ds' + q(s,r). \tag{4.1}$$

The specific intensity $u = u(s,r)$ depends on the position $r \in R$ and the direction of propagation described by a unit vector $s = (\cos\psi \sin\theta, \sin\psi \sin\theta, \cos\theta)^T$, $s \in S$, i.e., we assume a constant speed of propagation. The medium is characterized by the total attenuation coefficient $\sigma_t = \sigma_a + \sigma_s$, where σ_a and σ_s denote the absorption and scattering rates, respectively. The scattering phase function k relates pre- and post-collisional directions, and we consider exemplary the Henyey-Greenstein phase function

$$k(s \cdot s') = \frac{1}{4\pi} \frac{1-g^2}{[1-2g(s \cdot s') + g^2]^{3/2}}, \tag{4.2}$$

with anisotropy factor g. For $g = 0$, we speak about isotropic scattering, and for g close to one, we say that the scattering is (highly) forward peaked. For simplicity, we assume $0 \leq g < 1$ in the following. The case $-1 < g \leq 0$ is similar. Internal sources of radiation are modeled by the function q. Introducing the outer unit normal vector field $n(r)$ on ∂R, the boundary condition is modeled by

$$u(s,r) = f(s,r) \quad \text{for } (s,r) \in S \times \partial R \text{ such that } s \cdot n(r) < 0. \tag{4.3}$$

The content of this chapter was published in: J. Dölz, O. Palii and M. Schlottbom, Springer Journal of Scientific Computing, (2022) vol. 90, no. 94. https://doi.org/10.1007/s10915-021-01757-9

In this chapter we consider the iterative solution of the linear systems arising from the discretization of the anisotropic radiative transfer equations (4.1)–(4.3) by preconditioned Richardson iterations. We are particularly interested in robustly convergent methods for multiple physical regimes that, at the same time, can embody ballistic regimes $\sigma_s \ll 1$ and diffusive regimes, i.e., $\sigma_s \gg 1$ and $\sigma_a > 0$, and highly forward peaked scattering, as it occurs for example in medical imaging applications [26]. Due to the size of the arising systems of linear equations, their numerical solution is challenging, and a variety of methods were developed as briefly summarized next.

Related work

Since for realistic problems analytical solutions are not available, numerical approximations are required. Common discretization methods can be classified into two main approaches based on their semidiscretization in s. The spherical harmonics method [45, 21, 2] approximates the solution u by a truncated series of spherical harmonics, which allows for spectral convergence for smooth solutions. For non-smooth solutions, which is the generic situation, local approximations in s can be advantageous, which is achieved, e.g., by discrete ordinates methods [23, 2, 51, 22, 12], continuous Galerkin methods [8], the discontinuous Galerkin (DG) method [42, 3, 9], iteratively refined piecewise polynomial approximations [31], or hybrid methods [15, 19].

A common step in the solution of the linear systems resulting from local approximations in s is to split the discrete system into a transport part and a scattering part. While the inversion of transport is usually straight-forward, scattering introduces a dense coupling in s. The corresponding Richardson iteration resulting from this splitting is called the source iteration [6, 5], and it converges linearly with a rate $c = \|\sigma_s/\sigma_t\|_\infty$. For scattering dominated problems, such as the biomedical applications mentioned above, we have $c \approx 1$ and the convergence of the source iteration becomes too slow for such applications. Acceleration of the source iteration can be achieved by preconditioning, which usually employs the diffusion approximation to (4.1)–(4.3) [6], and the resulting scheme is then called diffusion synthetic accelerated (DSA) source iteration [6]. Although this approach is well motivated by asymptotic analysis, it faces several issues, such as, a proper generalization to multi-dimensional problems with anisotropy, strong variations in the optical parameters, or the use of unstructured and curved meshes, see [6].

Effective DSA schemes rely on consistent discretization of the corresponding diffusion approximation, see [9, 10] for isotropic scattering, and in [13] for two-dimensional problems with anisotropic scattering. The latter employs a modified interior penalty DG discretization for the corresponding diffusion approximation, which has also been used in [14] where it is, however, found that their DSA scheme becomes less effective for highly heterogeneous optical parameters. A discrete analysis of DSA schemes for high-order DG discretizations on possibly curved meshes, which may complicate the inversion of the transport part, can be found in [17]. In the variational framework of [9] consistency is automatically achieved by subspace correction instead of finding a consistent discretization of the diffusion approximation. This variational treatment allowed to prove convergence of the corresponding iteration and numerical re-

sults showed robust contraction rates, even in multi-dimensional calculations with heterogeneous optical parameters.

It is the purpose of this paper to generalize the approach of [9] to the anisotropic scattering case, which requires non-trivial extensions as outlined in the next section.

Approach and contribution

In this paper we focus on the construction of *robustly and provably convergent efficient iterative schemes* for the radiative transfer equation with anisotropic scattering. To describe our approach, let us introduce the linear system that we need to solve, which stems from a mixed finite element discretization of (4.1)–(4.3) using discontinuous polynomials on the sphere [7, 9], i.e.,

$$\begin{bmatrix} \mathbf{R}+\mathbf{M}^{+} & -\mathbf{A}^{\intercal} \\ \mathbf{A} & \mathbf{M}^{-} \end{bmatrix} \begin{bmatrix} \mathbf{u}^{+} \\ \mathbf{u}^{-} \end{bmatrix} = \begin{bmatrix} \mathbf{K}^{+} & \\ & \mathbf{K}^{-} \end{bmatrix} \begin{bmatrix} \mathbf{u}^{+} \\ \mathbf{u}^{-} \end{bmatrix} + \begin{bmatrix} \mathbf{q}^{+} \\ \mathbf{q}^{-} \end{bmatrix}. \tag{4.4}$$

Here, the superscripts in the equation refer to even ('+') and odd ('−') parts from the underlying discretization. The matrices $\mathbf{K}^{+}$ and $\mathbf{K}^{-}$ discretize scattering, while $\mathbf{R}$ incorporates boundary conditions, $\mathbf{M}^{+}$ and $\mathbf{M}^{-}$ are mass matrices related to σ_t, and $\mathbf{A}$ discretizes $s \cdot \nabla_r$, and their assembly can be done with standard FEM codes. The even part solves the even-parity equations

$$\mathbf{E}\mathbf{u}^{+} = \mathbf{K}^{+}\mathbf{u}^{+} + \mathbf{q}, \tag{4.5}$$

i.e., the Schur complement of (4.4), with symmetric positive definite matrix $\mathbf{E} = \mathbf{A}^{\intercal}(\mathbf{M}^{-}-\mathbf{K}^{-})^{-1}\mathbf{A}+\mathbf{M}^{+}+\mathbf{R}$ and source term $\mathbf{q} = \mathbf{q}^{+}+\mathbf{A}^{\intercal}(\mathbf{M}^{-}-\mathbf{K}^{-})^{-1}\mathbf{q}^{-}$. Once the even part $\mathbf{u}^{+}$ is known, the odd part $\mathbf{u}^{-}$ can be obtained from (4.4). The preconditioned Richardson iteration considered in this article then reads

$$\mathbf{u}^{+}_{n+1} = \big(\mathbf{I} - \mathbf{P}_2\mathbf{P}_1(\mathbf{E}-\mathbf{K}^{+})\big)\mathbf{u}^{+}_{n} + \mathbf{P}_2\mathbf{P}_1\mathbf{q}, \tag{4.6}$$

with preconditioners $\mathbf{P}_1$ and $\mathbf{P}_2$. Comparing to standard DSA source iterations, $\mathbf{P}_1$ corresponds to a transport sweep, and a typical choice that renders the convergence behavior of (4.6) independent of the discretization parameters is $\mathbf{P}_1 = \mathbf{E}^{-1}$. More precisely, we show that this choice of $\mathbf{P}_1$ yields a contraction rate of $c = \|\sigma_s/\sigma_t\|_\infty$. The second preconditioner $\mathbf{P}_2$ aims to improve the convergence behavior in diffusive regimes, $c \approx 1$. In the spirit of [9], we construct $\mathbf{P}_2$ via Galerkin projection onto suitable subspaces, which guarantees monotone convergence of (4.6). The construction of suitable subspaces that give good error reduction is motivated by the observation that error modes that are hardly damped by $\mathbf{I}-\mathbf{P}_1(\mathbf{E}-\mathbf{K}^{+})$ can be approximated well by spherical harmonics of low degree, cf. Section 4.3. While for the isotropic case $g = 0$, spherical harmonics of degree zero, i.e., constants in angle, are sufficient for obtaining good convergence rates, we show that higher order spherical harmonics should be used for anisotropic scattering. To preserve consistency, we replace higher order spherical harmonics, which are the eigenfunctions of the integral operator in (4.1), by discrete eigenfunctions of $\mathbf{K}^{+}$.

The efficiency of the proposed iterative scheme hinges on the ability to efficiently implement and apply the arising operators. While for $g = 0$, $\mathbf{K}^{-} = 0$, and $\mathbf{K}^{+}$ can be realized via fast Fourier transformation, and $\mathbf{E}$ is block-diagonal

with sparse blocks allowing for an efficient application of $\mathbf{E}$, the situation is more involved for $g > 0$. We show that $\mathbf{K}^+$ and $\mathbf{K}^-$ can be applied efficiently by exploiting their Kronecker structure between a sparse matrix and a dense matrix, which turns out to be efficiently applicable by using $\mathcal{H}$- or $\mathcal{H}^2$-matrix approximations independently of g. As we show the practical implementation of $\mathcal{H}$- or $\mathcal{H}^2$-matrices can be done by standard libraries, such as H2LIB [32] or BEMBEL [34]. This in combination with standard FEM assembly routines for the other matrices ensures robustness and maintainability of the code.

Since $\mathbf{A}$, $\mathbf{M}^+$, and $\mathbf{R}$ are sparse and block diagonal, the main bottleneck in the application of $\mathbf{E}$ is the application of $(\mathbf{M}^- - \mathbf{K}^-)^{-1}$. Based on the tensor structure of $\mathbf{K}^-$ and its spectral properties, we derive a preconditioner such that $(\mathbf{M}^- - \mathbf{K}^-)^{-1}$ can be applied robustly in g in only a few iterations. Thus, we can apply $\mathbf{E}$ in almost linear complexity. Efficiency of (4.6) is further increased by realizing the preconditioner $\mathbf{P}_1$ inexactly by employing a small, fixed number of l steps of an inner iterative scheme. Denoting the resulting preconditioner by $\mathbf{P}_1^l$, we show that the condition number of $\mathbf{P}_1^l\mathbf{E}$ is $O((1-(cg)^l)^{-1})$, which is robust in the limit $c \to 1$. In contrast, we note that the condition number of $\mathbf{P}_1^l(\mathbf{E}-\mathbf{K}^+)$ is $O((1-c)^{-1})$, i.e., a straight-forward iterative solution of the even-parity equations using a black-box solver, such as preconditioned conjugate gradients, is in general not robust for $c \to 1$.

Summarizing, each step of our iteration (4.6) can be performed very efficiently. The iteration is provably convergent and numerical results show that the contraction rates are robust for $c \to 1$. The result is a highly efficient numerical scheme for the solution of the even parity equations (4.5) and, thus, also for the overall system (4.4).

Outline

The structure of the chapter is as follows: In Section 4.2 we recall the variational formulation that builds the basis of our numerical scheme and establish some spectral equivalences for the scattering operator, which are key to the construction of our preconditioners. In Section 4.3 we present iterative schemes for the even-parity equations of radiative transfer in Hilbert space, which, after discretization in Section 4.4, result in the schemes described in Section 4.1. Details of the implementation and its complexity are described in Section 4.5. Numerical studies of the performance of the proposed methods and report on the results are presented in Section 4.7. The chapter closes with a discussion in Section 4.8.

4.2 Preliminaries

In the following we recall the relevant functional analytic framework, state the corresponding variational formulation of the radiative transfer problem (4.1)–(4.3) and provide some analytical results about the spectrum of the scattering operator, which we will later use for the construction of our preconditioners.

Function spaces

By $L^2(M)$ we denote the usual Hilbert space of square integrable functions on a manifold M, and denote $(u,w)_M = \int_M uw\,dM$ the corresponding inner product

and $\|u\|_{L^2(M)}$ the induced norm. For $M = D = S \times R$, we write $\mathbb{V} = L^2(D)$ and $(u, w) = (u, w)_D$. Functions $w \in \mathbb{V}$ with weak derivative $s \cdot \nabla_r w \in \mathbb{V}$ have a well-defined trace [11]. We restrict the natural trace space [11], and consider the weighted Hilbert space $L^2(\partial D_\pm; |s \cdot n|)$ of measurable functions w on

$$\partial D_\pm = \{(s, r) \in S \times \partial R : \pm s \cdot n(r) > 0\}$$

with $|s \cdot n|^{1/2} w \in L^2(\partial D_\pm)$. For the weak formulation of (4.1)–(4.3) we use the Hilbert space

$$\mathbb{W} = \{w \in L^2(D) : s \cdot \nabla_r w \in L^2(D),\ w_{|\partial D_-} \in L^2(\partial D_-; |s \cdot n|)\},$$

with corresponding norm $\|w\|^2_{\mathbb{W}} = \|s \cdot \nabla_r w\|^2_{L^2(D)} + \|w\|^2_{L^2(D)} + \|w\|^2_{L^2(\partial D_-; |s \cdot n|)}$.

Assumptions on the optical parameters and data

The data terms are assumed to satisfy $q \in L^2(D)$ and $f \in L^2(\partial D_-; |s \cdot n|)$. Absorption and scattering rates are non-negative and essentially bounded functions $\sigma_a, \sigma_s \in L^\infty(R)$. We assume that the medium occupied by R is absorbing, i.e., that there exists a constant $\gamma > 0$ such that $\sigma_a(r) \geq \gamma$ for a.e. $r \in R$. Thus, the ratio between the scattering rate and the total attenuation rate $\sigma_t = \sigma_a + \sigma_s$ is strictly less than one, $c = \|\sigma_s / \sigma_t\|_\infty < 1$.

Even-odd splitting

The space $\mathbb{V} = \mathbb{V}^+ \oplus \mathbb{V}^-$ allows for an orthogonal decomposition into even and odd functions of the variable $s \in S$. The even part u^+ and odd part u^- of a function $u \in \mathbb{V}$ is defined a.e. by

$$u^\pm(s, r) = \frac{1}{2}(u(s, r) \pm u(-s, r)).$$

Similarly, we denote $\mathbb{W}^\pm$ the corresponding subspaces of functions $u \in \mathbb{W}$ with $u \in \mathbb{V}^\pm$.

Operator formulation of the radiative transfer equation

The weak formulation of (4.1)–(4.3) presented in [7] can be stated concisely using suitable operators and we refer to [7] for proofs of the corresponding mapping properties. Let $u^+, w^+ \in \mathbb{W}^+$ and $u^- \in \mathbb{V}^-$. The transport operator $\mathcal{A} : \mathbb{W}^+ \to \mathbb{V}^-$ is defined by

$$\mathcal{A} u^+ = s \cdot \nabla_r u^+.$$

Identifying the dual $\mathbb{V}'$ of $\mathbb{V}$ with $\mathbb{V}$, the dual transport operator $\mathcal{A}' : \mathbb{V}^- \to (\mathbb{W}^+)'$ is defined by

$$\langle \mathcal{A}' u^-, w^+ \rangle = (\mathcal{A} w^+, u^-).$$

Boundary terms are handled by the operator $\mathcal{R} : \mathbb{W}^+ \to (\mathbb{W}^+)'$ defined by

$$\langle \mathcal{R} u^+, w^+ \rangle = (|s \cdot n| u^+, w^+)_{\partial D}.$$

Scattering is described by the operator $\mathcal{S} : L^2(S) \to L^2(S)$ defined by

$$(\mathcal{S}u)(s) = \int_S k(s \cdot s')u(s')ds',$$

where k is the phase function defined in (4.2). In slight abuse of notation, we also denote the trivial extension of $\mathcal{S}$ to an operator $L^2(D) \to L^2(D)$ by $\mathcal{S}$. We recall that $\mathcal{S}$ maps even to even and odd to odd functions [7, Lemma 2.6], and so does $\mathcal{K} : \mathbb{V} \to \mathbb{V}$ defined by

$$\mathcal{K}u = \sigma_s \mathcal{S}u.$$

We denote by $\mathcal{K}$ also its restrictions to $\mathbb{V}^{\pm}$ and $\mathbb{W}^+$, respectively. The spherical harmonics $\{H_m^l : l \in \mathbb{N} \cup \{0\}, -l \le m \le l\}$ form a complete orthogonal system for $L^2(S)$, and we assume the normalization $\|H_m^l\|_{L^2(S)} = 1$. Furthermore, H_m^l is an eigenfunction of $\mathcal{S}$ with eigenvalue g^l, i.e.,

$$\mathcal{S}H_m^l = g^l H_m^l, \tag{4.7}$$

and $H_m^l \in \mathbb{V}^+$ if l is an even number and $H_m^l \in \mathbb{V}^-$ if l is an odd number. Attenuation is described by the multiplication operator $\mathcal{M} : \mathbb{V} \to \mathbb{V}$ defined by

$$\mathcal{M}u = \sigma_t u.$$

Introducing the functionals $\ell^+ \in (\mathbb{W}^+)'$, and $\ell^- \in (\mathbb{V}^-)'$, given by

$$\begin{aligned} \ell^+(w^+) &= (q, w^+) + 2(|s \cdot n| f, w^+)_{\partial D_-}, && w^+ \in \mathbb{W}^+, \\ \ell^-(w^-) &= (q, w^-), && w^- \in \mathbb{V}^-, \end{aligned}$$

the operator formulation of the radiative transfer equation (4.1)–(4.3) is [7]: Find $(u^+, u^-) \in \mathbb{W}^+ \times \mathbb{V}^-$ such that

$$\mathcal{R}u^+ - \mathcal{A}'u^- + \mathcal{M}u^+ = \mathcal{K}u^+ + \ell^+ \qquad \text{in } (\mathbb{W}^+)', \tag{4.8}$$

$$\mathcal{A}u^+ + \mathcal{M}u^- = \mathcal{K}u^- + \ell^- \qquad \text{in } \mathbb{V}^-. \tag{4.9}$$

Well-posedness

In the situation of Section 4.2, there exists a unique solution $(u^+, u^-) \in \mathbb{W}^+ \times \mathbb{V}^-$ of (4.8) and (4.9) satisfying

$$\|u^+\|_{\mathbb{W}} + \|u^-\|_{\mathbb{V}} \le C(\|q\|_{L^2(D)} + \|f\|_{L^2(\partial D_-; |s \cdot n|)}),$$

with a constant C depending only on γ and $\|\sigma_t\|_\infty$ [7]. Notice that this well-posedness result remains true even if σ_a and σ_s are allowed to vanish [46]. As shown in [7, Theorem 4.1] it holds that $u^- \in \mathbb{W}^-$ and $u^+ + u^- \in \mathbb{W}$ satisfies (4.1) a.e. in D and (4.3) holds in $L^2(\partial D_-; |s \cdot n|)$.

Even-parity formulation

As in [7], it follows from (4.7) that for $v^- \in \mathbb{V}^-$

$$\inf_{r \in R} (\sigma_a + (1-g)\sigma_s)\|v^-\|_{\mathbb{V}}^2 \le \|v^-\|_{\mathcal{M}-\mathcal{K}}^2 \le \|\sigma_t\|_\infty \|v^-\|_{\mathbb{V}}^2, \tag{4.10}$$

where we write $\|w\|_{\mathcal{Q}}^2 = (\mathcal{Q}w, w)$ for any positive operator $\mathcal{Q}$.

Thus, $\mathcal{M} - \mathcal{K} : \mathbb{V}^- \to \mathbb{V}^-$ is boundedly invertible, and, by (4.9),

$$u^- = (\mathcal{M} - \mathcal{K})^{-1}(\ell^- - \mathcal{A}u^+). \tag{4.11}$$

Using (4.11) in (4.8) and introducing

$$\mathcal{E} : \mathbb{W}^+ \to (\mathbb{W}^+)', \quad \mathcal{E}u^+ = \mathcal{R}u^+ + \mathcal{A}'(\mathcal{M} - \mathcal{K})^{-1}\mathcal{A}u^+ + \mathcal{M}u^+,$$

and

$$\ell(w^+) = \ell^+(w^+) + ((\mathcal{M} - \mathcal{K})^{-1}q, \mathcal{A}w^+), \qquad w^+ \in \mathbb{W}^+,$$

the even-parity formulation of the radiative transfer equation is: Find $u^+ \in \mathbb{W}^+$ such that

$$(\mathcal{E} - \mathcal{K})u^+ = \ell. \tag{4.12}$$

As shown in [7], the even-parity formulation is a coercive, symmetric problem, which is well-posed by the Lax-Milgram lemma. Solving (4.12) for $u^+ \in \mathbb{W}^+$, we can retrieve $u^- \in \mathbb{V}^-$ by (4.11). In turn, $(u^+, u^-) \in \mathbb{W}^+ \times \mathbb{V}^-$ solves (4.8)–(4.9).

Preconditioning of $\mathcal{M} - \mathcal{K}$

We generalize the inequalities (4.10) to obtain spectrally equivalent approximations to $\mathcal{M} - \mathcal{K}$. Since $\mathcal{K} = \sigma_s\mathcal{S}$, we can construct approximations to $\mathcal{K}$ by approximating $\mathcal{S}$. To do so let us define for $N \in \mathbb{N}$ and $v \in \mathbb{V}$

$$\mathcal{S}_N v = \sum_{l=0}^{N} g^l \sum_{m=-l}^{l} (v, H_m^l)_S H_m^l. \tag{4.13}$$

Notice that the summation is only over even integers $0 \le l \le N$ if $v \in \mathbb{V}^+$ and only over odd ones if $v \in \mathbb{V}^-$. The approximation of $\mathcal{K}$ is then defined by $\mathcal{K}_N = \sigma_s\mathcal{S}_N$.

Lemma 4.2.1. *The operator $\mathcal{M} - \mathcal{K}_N$ is spectrally equivalent to $\mathcal{M} - \mathcal{K}$, that is*

$$\big(1 - cg^{N+1}\big)((\mathcal{M} - \mathcal{K}_N)v, v) \le ((\mathcal{M} - \mathcal{K})v, v) \le ((\mathcal{M} - \mathcal{K}_N)v, v)$$

for all $v \in \mathbb{V}$, with $c = \|\sigma_s/\sigma_t\|_\infty$. In particular, $\mathcal{M} - \mathcal{K}_N$ is invertible.

Proof. We use that $\{H_l^m\}$ is a complete orthonormal system of $L^2(S)$. Hence, any $v \in \mathbb{V} = L^2(S) \otimes L^2(R)$ has the expansion

$$v(s, r) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} v_m^l(r) H_m^l(s),$$

with $v_m^l \in L^2(R)$ and $\|v\|_{\mathbb{V}}^2 = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \|v_m^l\|_{L^2(R)}^2 < \infty$, and

$$((\mathcal{M} - \mathcal{K}_N)v, v) = \sum_{l=0}^{L} \sum_{m=-l}^{l} \left\| \sqrt{\sigma_t - g^l\sigma_s}\, v_m^l \right\|_{L^2(R)}^2 + \sum_{l=N+1}^{\infty} \sum_{m=-l}^{l} \left\| \sqrt{\sigma_t}\, v_m^l \right\|_{L^2(R)}^2.$$

Using $c = \|\sigma_s/\sigma_t\|_\infty$ it follows that

$$0 \le ((\mathcal{K} - \mathcal{K}_N)v, v) = \sum_{l=N+1}^{\infty} g^l \sum_{m=-l}^{l} \left\| \sqrt{\sigma_s} v_m^l \right\|_{L^2(R)}^2 \le c g^{N+1} ((\mathcal{M} - \mathcal{K}_N)v, v). \tag{4.14}$$

The inequalities in the statement then follow from

$$((\mathcal{M} - \mathcal{K})v, v) = ((\mathcal{M} - \mathcal{K}_N)v, v) - ((\mathcal{K} - \mathcal{K}_N)v, v),$$

while invertibility follows from [7, Lemma 2.14]. □

4.3 Iteration for the even-parity formulation

We generalize the Richardson iteration of [9] for the radiative transfer equation with isotropic scattering to the anisotropic case and equip the iteration process with a suitable preconditioner, which we will investigate later. We restrict ourselves to a presentation suitable for the error analysis and postpone the linear algebra setting and the discussion of its efficient realization to Section 4.5.

Derivation of the scheme

We consider the solution of (4.12) along the following two steps:
Step (i) Given $u_n^+ \in \mathbb{W}^+$ and a symmetric and positive definite operator $\mathcal{P}_1 : (\mathbb{W}^+)' \to \mathbb{W}^+$, we compute

$$u_{n+\frac{1}{2}}^+ = u_n^+ - \mathcal{P}_1((\mathcal{E} - \mathcal{K})u_n^+ - \ell). \tag{4.15}$$

Step (ii) Compute a subspace correction to $u_{n+1/2}^+$ based on the observation that the error $e_{n+1/2}^+ = u^+ - u_{n+1/2}^+$ satisfies

$$(\mathcal{E} - \mathcal{K})e_{n+\frac{1}{2}}^+ = ((\mathcal{E} - \mathcal{K})\mathcal{P}_1 - \mathcal{I})((\mathcal{E} - \mathcal{K})u_n^+ - \ell). \tag{4.16}$$

Solving (4.16) is as difficult as solving the original problem. Let $\mathbb{W}_N^+ \subset \mathbb{W}^+$ be closed, and consider the Galerkin projection $\mathcal{P}_G : \mathbb{W}^+ \to \mathbb{W}_N^+$ onto $\mathbb{W}_N^+$ defined by

$$\langle (\mathcal{E} - \mathcal{K})\mathcal{P}_G w, v\rangle = \langle (\mathcal{E} - \mathcal{K})w, v\rangle \quad \text{for all } v \in \mathbb{W}_N^+. \tag{4.17}$$

Using (4.16), the correction $u_{c,n}^+ = \mathcal{P}_G e_{n+1/2}^+$, is then characterized as the solution to

$$\langle (\mathcal{E} - \mathcal{K})u_{c,n}^+, v\rangle = \langle ((\mathcal{E} - \mathcal{K})\mathcal{P}_1 - \mathcal{I})((\mathcal{E} - \mathcal{K})u_n^+ - \ell), v\rangle \tag{4.18}$$

for all $v \in \mathbb{W}_N^+$, where the right-hand side involves available data only. The update is performed via

$$u_{n+1}^+ = u_{n+\frac{1}{2}}^+ + u_{c,n}^+. \tag{4.19}$$

Error analysis

Since $\mathcal{P}_G$ is non-expansive in the norm induced by $\mathcal{E}-\mathcal{K}$, the error analysis for the overall iteration (4.15) and (4.19) relies on the spectral properties of $\mathcal{P}_1$. Therefore, the following theoretical investigations consider the generalized eigenvalue problem

$$(\mathcal{E}-\mathcal{K})w = \lambda \mathcal{P}_1^{-1} w. \tag{4.20}$$

The following well-known lemma asserts that the half-step (4.15) yields a contraction if an appropriate preconditioner $\mathcal{P}_1$ is chosen. We provide a proof for later reference.

Lemma 4.3.1. *Let $0 < \beta \leq 1$ and assume that the eigenvalues λ of (4.20) satisfy $\beta \leq \lambda \leq 1$. Then, for any $u_n^+ \in \mathbb{W}^+$, $u_{n+1/2}^+$ defined via (4.15) satisfies*

$$\|u^+ - u_{n+\frac{1}{2}}^+\|_{\mathcal{E}-\mathcal{K}} \leq (1-\beta)\|u^+ - u_n^+\|_{\mathcal{E}-\mathcal{K}}.$$

Proof. Assume that $\{(w_k, \lambda_k)\}_{k\geq 0}$ is the eigensystem of the generalized eigenvalue problem (4.20). For any u_n^+, the error $e_n^+ = u^+ - u_n^+$ satisfies

$$e_{n+\frac{1}{2}}^+ = (\mathcal{I} - \mathcal{P}_1(\mathcal{E}-\mathcal{K}))e_n^+. \tag{4.21}$$

Using the expansion $e_n^+ = \sum_{k=0}^{\infty} a_k w_k$, we compute $\|e_n^+\|_{\mathcal{E}-\mathcal{K}}^2 = \sum_{k=0}^{\infty} a_k^2 \lambda_k$. Using (4.21), we thus obtain $e_{n+1/2}^+ = \sum_{k=0}^{\infty}(1-\lambda_k)a_k w_k$, and hence

$$\|e_{n+\frac{1}{2}}^+\|_{\mathcal{E}-\mathcal{K}}^2 = \sum_{k=0}^{\infty}(1-\lambda_k)^2\lambda_k a_k^2 \leq \sup_{0\leq k<\infty}(1-\lambda_k)^2 \|e_n^+\|_{\mathcal{E}-\mathcal{K}}^2.$$

Since $0 < \beta \leq \lambda_k \leq 1$ by assumption, the assertion follows. □

The next statement asserts that the iterative scheme defined by (4.19) converges linearly to the even part of the solution of the radiative transfer equation. It is a direct consequence of Lemma 4.3.1 and the observation that $e_{n+1}^+ = (\mathcal{I} - \mathcal{P}_G)e_{n+1/2}^+$ satisfies

$$\|e_{n+1}^+\|_{\mathcal{E}-\mathcal{K}} = \inf_{v\in\mathbb{W}_N^+} \|e_{n+\frac{1}{2}}^+ - v\|_{\mathcal{E}-\mathcal{K}}. \tag{4.22}$$

Lemma 4.3.2. *Let $\mathbb{W}_N^+ \subset \mathbb{W}^+$ be closed, and assume that the eigenvalues λ of (4.20) satisfy $\beta \leq \lambda \leq 1$ for some $0 < \beta \leq 1$. Then, for any $u_0^+ \in \mathbb{W}^+$, the sequence $\{u_n^+\}$ defined in (4.15) and (4.19) converges linearly to the solution u^+ of (4.12), i.e.,*

$$\|u^+ - u_{n+1}^+\|_{\mathcal{E}-\mathcal{K}} \leq (1-\beta)\|u^+ - u_n^+\|_{\mathcal{E}-\mathcal{K}}. \tag{4.23}$$

In view of the previous lemma fast convergence $u_n^+ \to u^+$ can be obtained by ensuring that β is close to one or by making the best-approximation error in (4.22) small. These two possibilities are discussed in the remainder of this section in more detail.

Generic preconditioners

The next result builds the basis for the preconditioner we will use later.

Lemma 4.3.3. *Let $\mathcal{P}_1$ be defined either by*
(i) $\mathcal{P}_1^{-1} = \mathcal{E}$ or
(ii) $\mathcal{P}_1^{-1} = \mathcal{E}_0 = (1-cg)^{-1}\mathcal{A}'\mathcal{M}^{-1}\mathcal{A} + \mathcal{M} + \mathcal{R}$.
Then $\mathcal{P}_1$ is spectrally equivalent to $\mathcal{E} - \mathcal{K}$, i.e.,

$$(1-c)(\mathcal{P}_1^{-1}w^+, w^+) \le ((\mathcal{E}-\mathcal{K})w^+, w^+) \le (\mathcal{P}_1^{-1}w^+, w^+),$$

for all $w^+ \in \mathbb{W}^+$. It holds $1-\beta = c$ in Lemma 4.3.2 in both cases.

Proof. Since $\mathcal{A}w^+ \in \mathbb{V}^-$, the result is a direct consequence of Lemma 4.2.1. □

Remark 4.3.4. *We can further generalize the choices for $\mathcal{P}_1^{-1}$ by choosing $N^+ \ge -1$, $N^- \ge 0$, and $\gamma_{N^-} = 1/(1-cg^{N^-+1})$. Then*

$$\mathcal{P}_1^{-1} = \mathcal{P}_{N^+,N^-}^{-1} = \mathcal{R} + \gamma_{N^-}\mathcal{A}'(\mathcal{M} - \mathcal{K}_{N^-})^{-1}\mathcal{A} + \mathcal{M} - \mathcal{K}_{N^+}$$

and $\mathcal{E} - \mathcal{K}$ are spectrally equivalent, i.e.,

$$(1-cg^{\min(N^-,N^+)+1})(\mathcal{P}_1^{-1}w^+, w^+) \le ((\mathcal{E}-\mathcal{K})w^+, w^+) \le (\mathcal{P}_1^{-1}w^+, w^+)$$

for all $w^+ \in \mathbb{W}^+$. In particular, $1-\beta = cg^{\min(N^-,N^+)+1}$ in Lemma 4.3.2.

Remark 4.3.5. *For isotropic scattering $g = 0$, we have that $\mathcal{E} = \mathcal{E}_0$. Thus, both choices in Lemma 4.3.3 can be understood as generalizations of the iteration considered in [9].*

The preconditioners in Remark 4.3.4 yield arbitrarily small contraction rates for sufficiently large N^+ and N^-. However, the efficient implementation of such a preconditioner seems to be rather challenging. Therefore, we focus on the preconditioners defined in Lemma 4.2.1 in the following. Since these choices for $\mathcal{P}_1$ yield slow convergence for $c \approx 1$, we need to construct $\mathbb{W}_N^+$ properly. This construction is motivated next, see Section 4.5 for a precise definition.

A motivation for constructing effective subspaces

From the proof of Lemma 4.3.1, one sees that error modes associated to small eigenvalues λ of (4.20) converge slowly. Hence, in order to regain fast convergence, such modes should be approximated well by functions in $\mathbb{W}_N^+$, see (4.22). Next, we give a heuristic motivation that such slowly convergent modes might be approximated well by low-order spherical harmonics.

Since we use $\mathcal{P}_1^{-1} \approx \mathcal{E}$ below, let us fix $\mathcal{P}_1^{-1} = \mathcal{E}$ in this subsection. Furthermore, let w be a slowly damped mode, i.e., w satisfies (4.20) with λ such that $\lambda \approx 1-c \approx 0$. Observe that w also satisfies $\mathcal{K}w = \delta\mathcal{E}w$ with $\delta = 1-\lambda \approx c \approx 1$, and $\delta \le c$ by Lemma 4.3.3(i). Let us expand the angular part of w into spherical harmonics, cf. Section 4.2,

$$w(s,r) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} w_m^l(r)H_m^l(s),$$

where $w_m^l = 0$ if l is odd. As in the proof of lemma 4.3.3, we obtain

$$\mathcal{K}w = \sum_{l=0}^{\infty} g^l \sum_{m=-l}^{l} \sigma_s(r) w_m^l(r) H_m^l(s).$$

Since $\sigma_s \leq \sigma_t$, orthogonality of the spherical harmonics implies

$$\sum_{l=0}^{\infty} cg^l \sum_{m=-l}^{l} \|\sqrt{\sigma_t} w_m^l\|_{L^2(R)}^2 \geq (\mathcal{K}w, w) =$$

$$\delta\left(\langle \mathcal{R}w, w\rangle + \|s\cdot\nabla_r w\|_{(\mathcal{M}-\mathcal{K})^{-1}}^2 + \sum_{l=0}^{\infty}\sum_{m=-l}^{l} \|\sqrt{\sigma_t} w_m^l\|_{L^2(R)}^2 \right).$$

Neglecting the contributions from $\mathcal{R}$ and $s\cdot\nabla_r$, we see that

$$\sum_{l=0}^{\infty} (cg^l - \delta) \sum_{m=-l}^{l} \|\sqrt{\sigma_t} w_m^l\|_{L^2(R)}^2 \geq 0. \tag{4.24}$$

Since $\delta \approx c \approx 1$ by assumption and $g < 1$, (4.24) can hold true only if w can be approximated well by spherical harmonics of degree less than or equal to N for some moderate integer N.

To convince the reader that this is likely to be true, we consider in the following the case $g = 0$ and remark that the overall behaviour does not change too much when varying g. If $c = \delta$, then (4.24) implies that $w_m^l = 0$ for all $l > 0$. If $\delta < c$, then (4.24) is equivalent to

$$\|\sqrt{\sigma_t} w_0^0\|_{L^2(R)}^2 \geq \frac{\delta}{c-\delta} \sum_{l=1}^{\infty}\sum_{m=-l}^{l} \|\sqrt{\sigma_t} w_m^l\|_{L^2(R)}^2.$$

Therefore, using orthogonality of the spherical harmonics once more, we obtain

$$\sum_{l=1}^{\infty}\sum_{m=-l}^{l} \|\sqrt{\sigma_t} w_m^l\|_{L^2(R)}^2 = \|\sqrt{\sigma_t} w\|_{L^2(D)}^2 - \|\sqrt{\sigma_t} w_0^0\|_{L^2(R)}^2$$

$$\leq \|\sqrt{\sigma_t} w\|_{L^2(D)}^2 - \frac{\delta}{c-\delta} \sum_{l=1}^{\infty}\sum_{m=-l}^{l} \|\sqrt{\sigma_t} w_m^l\|_{L^2(R)}^2.$$

Rearranging terms yields the estimate

$$\sum_{l=1}^{\infty}\sum_{m=-l}^{l} \|\sqrt{\sigma_t} w_m^l\|_{L^2(R)}^2 \leq \big(1 - \delta/c\big) \|\sqrt{\sigma_t} w\|_{L^2(D)}^2.$$

Since, by assumption, $\delta \approx c$, we conclude that w can be approximated well by $w_0^0 H_0^0$.

Note that this statement quantifies approximation in terms of the L^2-norm. However, using recurrence relations of spherical harmonics to incorporate the terms $\langle \mathcal{R}w, w\rangle + \|s\cdot\nabla_r w\|_{(\mathcal{M}-\mathcal{K})^{-1}}^2$ into (4.24), suggests that a similar statement also holds for the $\mathcal{E}-\mathcal{K}$-norm. A full analysis of this statement seems out of the scope of this paper, and we postpone it to future research. We conclude that effective subspaces $\mathbb{W}_N^+$ consist of linear combinations of low-order spherical harmonics, and we employ this observation in our numerical realization.

4.4 Galerkin approximation

The iterative scheme of the previous section has been formulated for infinite-dimensional function spaces $\mathbb{W}^+$ and $\mathbb{W}_N^+ \subset \mathbb{W}^+$. For the practical implementation we recall the approximation spaces described in [7] and [9, Section 6.3]. Let $\mathcal{T}_h^R$ and $\mathcal{T}_h^S$ denote shape regular triangulations of R and S, respectively. For simplicity we assume the triangulations to be quasi-uniform. To properly define even and odd functions associated with the triangulations, we further require that $-K_S \in \mathcal{T}_h^S$ for each spherical element $K_S \in \mathcal{T}_h^S$. The latter requirement can be ensured by starting with a triangulation of a half-sphere and reflection. Let $\mathbb{X}_h^+ = \mathbb{P}_1^c(\mathcal{T}_h^R)$ denote the vector space of continuous, piecewise linear functions subordinate to the triangulation $\mathcal{T}_h^R$ with basis $\{\varphi_i\}$ and dimension n_R^+, and let $\mathbb{X}_h^- = \mathbb{P}_0(\mathcal{T}_h^R)$ denote the vector space of piecewise constant functions subordinate to $\mathcal{T}_h^R$ with basis $\{\chi_j\}$ and dimension n_R^-. Similarly, we denote by $\mathbb{S}_h^+ = \mathbb{P}_0(\mathcal{T}_h^S) \cap L^2(S)^+$ and $\mathbb{S}_h^- = \mathbb{P}_1(\mathcal{T}_h^S) \cap L^2(S)^-$ the vector spaces of even, piecewise constant and odd, piecewise linear functions subordinate to the triangulation $\mathcal{T}_h^S$, respectively. We can construct a basis $\{\mu_k^+\}$ for $\mathbb{S}_h^+$ by choosing n_S^+ many triangles with midpoints in a given half-sphere, and define the functions μ_k^+ to be the indicator functions of these triangles. For any other point $s \in S$, we find $K_S \in \mathcal{T}_h^S$ with midpoint in the given half-sphere such that $-s \in K_S$ and we define $\mu_k^+(s) = \mu_k^+(-s)$. A similar construction leads to a basis $\{\psi_l^-\}$ of $\mathbb{S}_h^-$. The conforming approximation spaces are then defined through tensor product constructions, $\mathbb{W}_h^+ = \mathbb{S}_h^+ \otimes \mathbb{X}_h^+$, $\mathbb{V}_h^- = \mathbb{S}_h^- \otimes \mathbb{X}_h^-$. Thus, for some coefficient matrices $\left[\mathbf{U}_{i,k}^+\right] \in \mathbb{R}^{n_R^+ \times n_S^+}$ and $\left[\mathbf{U}_{j,l}^-\right] \in \mathbb{R}^{n_R^- \times n_S^-}$, any $u_h^+ \in \mathbb{W}_h^+$ and $u_h^- \in \mathbb{V}_h^-$ can be expanded as

$$u_h^+ = \sum_{i=1}^{n_R^+} \sum_{k=1}^{n_S^+} \mathbf{U}_{i,k}^+ \varphi_i \mu_k^+, \qquad u_h^- = \sum_{j=1}^{n_R^-} \sum_{l=1}^{n_S^-} \mathbf{U}_{j,l}^- \chi_j \psi_l^-. \tag{4.25}$$

The Galerkin approximation of (4.8)–(4.9) computes $(u_h^+, u_h^-) \in \mathbb{W}_h^+ \times \mathbb{V}_h^-$ such that

$$\mathcal{R} u_h^+ - \mathcal{A}' u_h^- + \mathcal{M} u_h^+ = \mathcal{K} u_h^+ + \ell^+ \qquad \text{in } (\mathbb{W}_h^+)', \tag{4.26}$$

$$\mathcal{A} u_h^+ + \mathcal{M} u_h^- = \mathcal{K} u_h^- + \ell^- \qquad \text{in } \mathbb{V}_h^-. \tag{4.27}$$

The discrete mixed system (4.26)–(4.27) can be solved uniquely [7]. Denoting $\mathbf{u}^\pm = \mathrm{vec}(\mathbf{U}^\pm)$ the concatenation of the columns of the matrices $\mathbf{U}^\pm$ into a vector, the mixed system (4.26)–(4.27) can be written as the following linear system

$$\begin{bmatrix} \mathbf{R} + \mathbf{M}^+ & -\mathbf{A}^\intercal \\ \mathbf{A} & \mathbf{M}^- \end{bmatrix} \begin{bmatrix} \mathbf{u}^+ \\ \mathbf{u}^- \end{bmatrix} = \begin{bmatrix} \mathbf{K}^+ & \\ & \mathbf{K}^- \end{bmatrix} \begin{bmatrix} \mathbf{u}^+ \\ \mathbf{u}^- \end{bmatrix} + \begin{bmatrix} \mathbf{q}^+ \\ \mathbf{q}^- \end{bmatrix}. \tag{4.28}$$

The matrices in the system are given by

$$\mathbf{K}^+ = \mathbf{S}^+ \otimes \mathfrak{M}_s^+, \qquad \mathbf{K}^- = \mathbf{S}^- \otimes \mathfrak{M}_s^-, \tag{4.29}$$

$$\mathbf{M}^+ = \mathbf{M}^+ \otimes \mathfrak{M}_t^+, \qquad \mathbf{M}^- = \mathbf{M}^- \otimes \mathfrak{M}_t^-, \tag{4.30}$$

$$\mathbf{A} = \sum_{i=1}^{d} \mathbf{A}_i \otimes \mathfrak{D}_i, \qquad \mathbf{R} = \mathrm{blkdiag}(\mathfrak{R}_1, \ldots, \mathfrak{R}_{n_S^+}), \tag{4.31}$$

where we denote by Gothic letters the matrices arising from the discretization on R and by Sans Serif letters matrices arising from the discretization on S, i.e.,

$$
\begin{aligned}
(\mathfrak{M}_t^-)_{j,j'} &= \int_R \sigma_t \chi_j \chi_{j'} dr, & (\mathsf{S}^-)_{l,l'} &= \int_S \mathcal{S}\psi_l^- \psi_{l'}^- ds,\\
(\mathfrak{M}_t^+)_{i,i'} &= \int_R \sigma_t \varphi_i \varphi_{i'} dr, & (\mathsf{S}^+)_{k,k'} &= \int_S \mathcal{S}\mu_k^+ \mu_{k'}^+ ds,\\
(\mathfrak{D}_n)_{j,i} &= \int_R \frac{\partial \varphi_i}{\partial r_n} \chi_j dr, & (\mathsf{A}_n)_{l,k} &= \int_S s_n \psi_l^- \mu_k^+ ds,\\
(\mathfrak{R}_k)_{i,i'} &= \int_{\partial R} \varphi_i \varphi_{i'} \omega_k dr, & \omega_k &= \int_S |s\cdot n|(\mu_k^+)^2 ds.
\end{aligned}
$$

The matrices $\mathfrak{M}_s^-$ and $\mathfrak{M}_s^+$ are defined accordingly. By M^+ and M^- we denote the Gramian matrices in $L^2(S)$.

We readily remark that all of these matrices are sparse, except for S^+ and S^-, which are dense. M^+ and M^- are diagonal and 3×3 block diagonal, respectively. Moreover, we note that $\mathfrak{M}_t^-$ is a diagonal matrix.

To conclude this section let us remark that taking the Schur complement of (4.28) finally yields the matrix counterpart of the even-parity system (4.12), i.e.,

$$\mathbf{E}\mathbf{u}^+ = \mathbf{K}^+\mathbf{u}^+ + \mathbf{q} \tag{4.32}$$

with $\mathbf{E} = \mathbf{A}^\intercal(\mathbf{M}^- - \mathbf{K}^-)^{-1}\mathbf{A} + \mathbf{M}^+ + \mathbf{R}$ and $\mathbf{q} = \mathbf{q}^+ + \mathbf{A}^\intercal(\mathbf{M}^- - \mathbf{K}^-)^{-1}\mathbf{q}^-$.

4.5 Discrete preconditioned Richardson iteration

After discretization, the iteration presented in Section 4.3 becomes

$$\mathbf{u}_{n+1}^+ = \mathbf{u}_n^+ - \mathbf{P}_2\mathbf{P}_1((\mathbf{E}-\mathbf{K}^+)\mathbf{u}_n^+ - \mathbf{q}). \tag{4.33}$$

The preconditioner $\mathbf{P}_1$ is directly related to $\mathcal{P}_1$ in (4.15). By denoting the coordinate vectors of the basis functions of the subspace $\mathbb{W}_{h,N}^+ \subset \mathbb{W}_h^+$ by $\mathbf{W}$, the matrix representation of the overall preconditioner is

$$\mathbf{P}_2\mathbf{P}_1 = \mathbf{P}_1 + \mathbf{W}\big(\mathbf{W}^\intercal(\mathbf{E}-\mathbf{K}^+)\mathbf{W}\big)^{-1}\mathbf{W}^\intercal(\mathbf{I}^+ - (\mathbf{E}-\mathbf{K}^+)\mathbf{P}_1). \tag{4.34}$$

Denoting $\mathbf{P}_G = \mathbf{W}\big(\mathbf{W}^\intercal(\mathbf{E}-\mathbf{K}^+)\mathbf{W}\big)^{-1}\mathbf{W}^\intercal(\mathbf{E}-\mathbf{K}^+)$ the matrix representation of the Galerkin projection $\mathcal{P}_G$ defined in (4.17), the iteration matrix admits the factorization

$$\mathbf{I}^+ - \mathbf{P}_2\mathbf{P}_1(\mathbf{E}-\mathbf{K}^+) = (\mathbf{I}^+ - \mathbf{P}_G)\big(\mathbf{I}^+ - \mathbf{P}_1(\mathbf{E}-\mathbf{K}^+)\big).$$

The discrete analog of Lemma 4.3.2 implies that the sequence $\{\mathbf{u}_n^+\}$ generated by (4.33) converges for any initial choice $\mathbf{u}_0^+$ to the solution $\mathbf{u}^+$ of (4.32). More precisely, by choosing $\mathbf{P}_1$ according to Lemma 4.3.3, there holds

$$\|\mathbf{u}^+ - \mathbf{u}_{n+1}^+\|_{\mathbf{E}-\mathbf{K}^+} \le \eta\|\mathbf{u}^+ - \mathbf{u}_n^+\|_{\mathbf{E}-\mathbf{K}^+}, \tag{4.35}$$

where $0 \le \eta \le c < 1$ is defined as

$$\eta = \sup \|(\mathbf{I}^+ - \mathbf{P}_G)(\mathbf{I}^+ - \mathbf{P}_1(\mathbf{E}-\mathbf{K}^+))\mathbf{v}^+\|_{\mathbf{E}-\mathbf{K}^+} \tag{4.36}$$

with supremum taken over all $\mathbf{v}^+ \in \mathbb{R}^{n_S^+ n_R^+}$ satisfying $\|\mathbf{v}^+\|_{\mathbf{E}-\mathbf{K}^+} = 1$. The realization of (4.33) relies on the efficient application of $\mathbf{E}$, $\mathbf{K}^+$, $\mathbf{P}_1$ and $\mathbf{P}_2$ discussed next.

Application of E

In view of (4.30) and (4.31) it is clear that $\mathbf{A}$, $\mathbf{M}^+$, and $\mathbf{M}^-$ can be stored and applied efficiently by using their tensor product structure, sparsity, and the characterization

$$(\mathbf{B} \otimes \mathbf{C})\,\mathrm{vec}(\mathbf{X}) = \mathrm{vec}(\mathbf{D}) \quad \Longleftrightarrow \quad \mathbf{C}\mathbf{X}\mathbf{B}^\intercal = \mathbf{D}, \tag{4.37}$$

where $\mathbf{C} \in \mathbb{R}^{m\times n}$, $\mathbf{X} \in \mathbb{R}^{n\times p}$, $\mathbf{B} \in \mathbb{R}^{q\times p}$, $\mathbf{D} \in \mathbb{R}^{m\times q}$. The boundary matrix $\mathbf{R}$ consists of sparse diagonal blocks, and can thus also be applied efficiently, see Section 4.6 for details. The remaining operation required for the application of $\mathbf{E}$ as given in (4.32) is the application of $(\mathbf{M}^- - \mathbf{K}^-)^{-1}$, which deserves some discussion. Since $\mathbf{M}^- - \mathbf{K}^-$ has a condition number of $(1-cg)^{-1}$ due to Lemma 4.2.1, a straightforward implementation with the conjugate gradient method may be inefficient for $cg \approx 1$.

To mitigate the influence of cg, we can use Lemma 4.2.1 once more and obtain preconditioners derived from $\mathcal{M} - \mathcal{K}_N$, which lead to bounds on the condition number by $(1-(cg)^{N+2})^{-1}$ for odd N. In what follows, we comment on the practical realization of such preconditioners and their numerical construction. As we will verify in the numerical examples, these preconditioners allow the application of $(\mathbf{M}^- - \mathbf{K}^-)^{-1}$ in only a few iterations even for g close to 1.

After discretization, the continuous eigenvalue problem (4.7) for the scattering operator becomes the generalized eigenvalue problem

$$\mathbf{S}^-\mathbf{W}^- = \mathbf{M}^-\mathbf{W}^-\mathbf{\Lambda}^-.$$

Since $\mathbf{S}^-$ and $\mathbf{M}^-$ are symmetric and positive, the eigenvalues satisfy $0 \le \lambda_l \le g$, and we assume that they are ordered non-increasingly. The eigenvectors $\mathbf{W}^-$ form an orthonormal basis $(\mathbf{W}^-)^\intercal\mathbf{M}^-\mathbf{W}^- = \mathbf{I}^-$. Truncation of the eigendecomposition at index $d_N = (N+1)(N+2)/2$, N odd, which is the number of odd spherical harmonics of order less than or equal to N, yields the approximation

$$\mathbf{S}^- = \mathbf{M}^-\mathbf{W}^-\mathbf{\Lambda}^-(\mathbf{W}^-)^\intercal\mathbf{M}^- \approx \mathbf{M}^-\mathbf{W}_N^-\mathbf{\Lambda}_N^-(\mathbf{W}_N^-)^\intercal\mathbf{M}^- =: \mathbf{S}_N^-. \tag{4.38}$$

The discrete version of $\mathcal{M} - \mathcal{K}_N$ then reads $\mathbf{M}^- - \mathbf{K}_N^-$, with $\mathbf{K}_N^- = \mathbf{S}_N^- \otimes \mathfrak{M}_s^-$. An explicit representation of its inverse is given by the following lemma. Its essential idea is to use an orthogonal decomposition of $\mathbb{V}_h^-$ induced by the eigen decomposition of $\mathbf{S}^-$, and to employ the diagonal representation of $\mathbf{M}^- - \mathbf{K}_N^-$ in the angular eigenbasis.

Lemma 4.5.1. *Let* $\mathbf{b} \in \mathbb{R}^{n_S^- n_R^-}$. *Then* $\mathbf{x} = (\mathbf{M}^- - \mathbf{K}_N^-)^{-1}\mathbf{b}$ *is given by*

$$\begin{aligned}\mathbf{x} = \Big(\mathbf{W}_N^- \otimes \mathfrak{I}^-\Big)\Big(\mathbf{I}^- \otimes \mathfrak{M}_t^- - \mathbf{\Lambda}_N^- \otimes \mathfrak{M}_s^-\Big)^{-1}\Big((\mathbf{W}_N^-)^\intercal \otimes \mathfrak{I}^-\Big)\mathbf{b} \\ + \Big(\Big((\mathbf{M}^-)^{-1} - \mathbf{W}_N^-(\mathbf{W}_N^-)^\intercal\Big) \otimes (\mathfrak{M}_t^-)^{-1}\Big)\mathbf{b},\end{aligned} \tag{4.39}$$

where $\mathfrak{I}^-$ *and* $\mathbf{I}^-$ *denote the identity matrices of dimension* n_R^- *and* d_N, *respectively.*

Proof. We first decompose $\mathbf{x}$ as follows

$$\mathbf{x} = (\mathbf{W}_N^-(\mathbf{W}_N^-)^\intercal \mathbf{M}^- \otimes \mathfrak{I}^-)\mathbf{x} + \big((\mathbf{I}^- - \mathbf{W}_N^-(\mathbf{W}_N^-)^\intercal \mathbf{M}^-) \otimes \mathfrak{I}^-\big)\mathbf{x}. \tag{4.40}$$

Applying $(\mathbf{W}_N^-)^\intercal \otimes \mathfrak{I}^-$ to $(\mathbf{M}^- - \mathbf{K}_N^-)\mathbf{x} = \mathbf{b}$, (4.38), and $\mathbf{M}^-$-orthogonality of $\mathbf{W}_N^-$ yield

$$\big(\mathbf{I}^- \otimes \mathfrak{M}_t^- - \mathbf{\Lambda}_N^- \otimes \mathfrak{M}_s^-\big)((\mathbf{W}_N^-)^\intercal \mathbf{M}^- \otimes \mathfrak{I}^-)\mathbf{x} = ((\mathbf{W}_N^-)^\intercal \otimes \mathfrak{I}^-)\mathbf{b}.$$

Inverting $\mathbf{I}^- \otimes \mathfrak{M}_t^- - \mathbf{\Lambda}_N^- \otimes \mathfrak{M}_s^-$ and applying $\mathbf{W}_N^- \otimes \mathfrak{I}^-$ further yields

$$(\mathbf{W}_N^-(\mathbf{W}_N^-)^\intercal \mathbf{M}^- \otimes \mathfrak{I}^-)\mathbf{x} = (\mathbf{W}_N^- \otimes \mathfrak{I}^-)\big(\mathbf{I}^- \otimes \mathfrak{M}_t^- - \mathbf{\Lambda}_N^- \otimes \mathfrak{M}_s^-\big)^{-1}((\mathbf{W}_N^-)^\intercal \otimes \mathfrak{I}^-)\mathbf{b}.$$

For the other part in (4.40), apply $((\mathbf{M}^-)^{-1} - \mathbf{W}_N^-(\mathbf{W}_N^-)^\intercal) \otimes (\mathfrak{M}_t^-)^{-1}$ to $(\mathbf{M}^- - \mathbf{K}_N^-)\mathbf{x} = \mathbf{b}$ and obtain

$$((\mathbf{I}^- - \mathbf{W}_N^-(\mathbf{W}_N^-)^\intercal \mathbf{M}^-) \otimes \mathfrak{I}^-)\mathbf{x} = \big(((\mathbf{M}^-)^{-1} - \mathbf{W}_N^-(\mathbf{W}_N^-)^\intercal) \otimes (\mathfrak{M}_t^-)^{-1}\big)\mathbf{b}.$$

Substituting both expressions into (4.40) yields the assertion. □

Remark 4.5.2. *If σ_s has huge variations, a more effective approximation to $\mathbf{K}^-$ can be obtained from the eigendecomposition*

$$\mathfrak{M}_s^- \mathfrak{I}^- = \mathfrak{M}_t^- \mathfrak{I}^- \Delta$$

with diagonal matrix Δ with entries $\Delta_j = \int_R \sigma_s \chi_j dr / \int_R \sigma_t \chi_j dr$. The modified approximation $\widetilde{\mathbf{K}^-}$ is then computed by considering only those combinations of spatial and angular eigenfunctions for which $\lambda_l \Delta_j$ is above a certain tolerance.

Application of $\mathbf{K}^+$ and $\mathbf{K}^-$

Although $\mathbf{K}^+$ and $\mathbf{K}^-$ provide a tensor product structure (4.29) involving the sparse matrices $\mathfrak{M}_s^+$ and $\mathfrak{M}_s^-$, the density of the scattering operators $\mathbf{S}^+$ and $\mathbf{S}^-$ becomes a bottleneck for iterative methods due to quadratic complexity in storage consumption and computational cost for assembly and matrix-vector products. $\mathcal{H}$- and $\mathcal{H}^2$-matrices, which can be considered as abstract variants of the fast multipole method [35, 39], where developed in the context of the boundary element method and can realize the storage, assembly and matrix-vector multiplication in linear or almost linear complexity, see [41, 40] and the references therein. A sufficient condition for compressibility in these formats is the following.

Definition 1. *Let $\tilde{S} \subset \mathbb{R}^d$ such that $k\colon \tilde{S} \times \tilde{S} \to \mathbb{R}$ is defined and arbitrarily often differentiable for all $\tilde{\mathbf{x}} \neq \tilde{\mathbf{y}}$ with $\tilde{\mathbf{x}}, \tilde{\mathbf{y}} \in \tilde{S}$. Then k is called* asymptotically smooth *if*

$$\big|\partial_{\tilde{\mathbf{x}}}^{\boldsymbol{\alpha}} \partial_{\tilde{\mathbf{y}}}^{\boldsymbol{\beta}} k(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})\big| \leq C \frac{(|\boldsymbol{\alpha}| + |\boldsymbol{\beta}|)!}{r^{|\boldsymbol{\alpha}|+|\boldsymbol{\beta}|}} \|\tilde{\mathbf{x}} - \tilde{\mathbf{y}}\|^{-|\boldsymbol{\alpha}|-|\boldsymbol{\beta}|}, \qquad \tilde{\mathbf{x}} \neq \tilde{\mathbf{y}}, \tag{4.41}$$

independently of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ for some constants $C, r > 0$.

While several methods [36, 37] can operate on the Henyey-Greenstein kernel on the sphere, most classical methods require an extension into space which we define as

$$K(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) = k(\mathbf{x} \cdot \mathbf{y}), \qquad \text{with } \mathbf{x} = \tilde{\mathbf{x}}/\|\tilde{\mathbf{x}}\|,\ \mathbf{y} = \tilde{\mathbf{y}}/\|\tilde{\mathbf{y}}\|. \tag{4.42}$$

The following result allows to use this extension in most $\mathcal{H}$- and $\mathcal{H}^2$-matrix libraries such as [32, 34, 33] in a black-box fashion.

Lemma 4.5.3. *Let $g \geq 0$. Then $K(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$ is asymptotically smooth for $\tilde{\mathbf{x}}, \tilde{\mathbf{y}} \in \mathbb{R}^d \setminus \{0\}$.*

Proof. We first remark that the cosinus theorem implies for $\mathbf{x}, \mathbf{y} \in S$ with angle φ that $\mathbf{x} \cdot \mathbf{y} = \cos(\varphi) = 1 - \|\mathbf{x} - \mathbf{y}\|^2/2$. Moreover, $\tilde{k}(\xi) = k(1 - \xi^2/2)$ is holomorphic for $\Re(\xi) > 0$ such that its Taylor series around $\xi > 0$ has convergence radius ξ and the derivatives of $\tilde{k}$ satisfy $\left|\partial_\xi^\alpha \tilde{k}(\xi)\right| \leq c r^\alpha \alpha! |\xi|^{-\alpha}$, $\alpha \in \mathbb{N}_0$, for all $\xi > 0$. Since $\tilde{\mathbf{x}} \mapsto \mathbf{x} = \tilde{\mathbf{x}}/\|\tilde{\mathbf{x}}\|$ is analytic for $\tilde{\mathbf{x}} \neq 0$ and since $K(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) = \tilde{k}(\|\mathbf{x} - \mathbf{y}\|)$, the assertion follows in complete analogy to the appendix of [38]. □

The $\mathcal{H}$- or $\mathcal{H}^2$-approximation of $\mathbf{S}^+$ and $\mathbf{S}^-$ and the sparsity of $\mathfrak{M}_s^+$ and $\mathfrak{M}_s^-$ combined with the tensor product identity (4.37) then allow for an application of $\mathbf{K}^+$ and $\mathbf{K}^-$ in almost linear or even linear complexity.

Choice and implementation of $\mathbf{P}_1$

As shown in Section 4.3, choosing $\mathbf{P}_1$ as in Lemma 4.3.3 leads to contraction rates $\eta \leq c$ in (4.35), i.e., independent of the mesh-parameters. The choice $\mathbf{P}_1 = \mathbf{E}^{-1}$ can be realized through an inner iterative methods, such as a preconditioned Richardson iteration resulting in an inner-outer iteration scheme when employed in (4.33). An effective preconditioner for $\mathbf{E}$ is given by the block-diagonal, symmetric positive definite matrix $\mathbf{E}_0 = \frac{1}{1-cg}\mathbf{A}^\intercal(\mathbf{M}^-)^{-1}\mathbf{A} + \mathbf{R} + \mathbf{M}^+$ which provides the spectral estimates

$$(1 - cg)\mathbf{x}^\intercal \mathbf{E}_0 \mathbf{x} \leq \mathbf{x}^\intercal \mathbf{E} \mathbf{x} \leq \mathbf{x}^\intercal \mathbf{E}_0 \mathbf{x}, \tag{4.43}$$

for all $\mathbf{x} \in \mathbb{R}^{n_S^+ n_R^+}$, cf. Lemma 4.2.1. Thus, the condition number of $\mathbf{E}_0^{-1}\mathbf{E}$ is bounded by $(1 - cg)^{-1}$, which is uniformly bounded for $c \in [0, 1]$ for fixed $g < 1$. For clarity of presentation, we will use a preconditioned Richardson iteration for the inner iteration to implement $\mathbf{P}_1$ in the rest of the paper, but remark that a non-stationary preconditioned conjugate gradient method will lead to even better performance. Applying $\mathbf{P}_1$ with high accuracy may still involve many iterations. Instead, we use a preconditioner $\mathbf{P}_1^l$ which performs l steps of an inner iteration, i.e., we set $\mathbf{P}_1^l \mathbf{b} = \mathbf{z}_l$, where

$$\mathbf{z}_0 = 0, \qquad \mathbf{z}_{k+1} = \mathbf{z}_k - \mathbf{E}_0^{-1}(\mathbf{E}\mathbf{z}_k - \mathbf{b}), \quad k < l. \tag{4.44}$$

Notice that, $\mathbf{P}_1^1 = \mathbf{E}_0^{-1}$ while $\mathbf{P}_1^l b \to \mathbf{E}^{-1}\mathbf{b}$ as $l \to \infty$. In fact, with similar arguments as in Lemma 4.3.1, it follows from (4.43) that

$$\|\mathbf{P}_1^l \mathbf{b} - \mathbf{E}^{-1}\mathbf{b}\|_{\mathbf{E}} \leq (cg)^l \|\mathbf{E}^{-1}\mathbf{b}\|_{\mathbf{E}}, \tag{4.45}$$

where $\|\mathbf{x}\|_{\mathbf{E}}^2 = \mathbf{x}^\intercal \mathbf{E} \mathbf{x}$. The next result asserts that this inexact realization of the preconditioner leads to a convergent scheme.

Lemma 4.5.4. *Let $l \geq 1$ be fixed. The iteration* (4.32) *with preconditioner $\mathbf{P}_1 = \mathbf{P}_1^l$ defines a convergent sequence, i.e.,* (4.35) *holds with $\eta \leq c$ and η as in* (4.36)*.*

Proof. Observing that $\mathbf{P}_1^l = \sum_{k=0}^{l-1}(\mathbf{E}_0^{-1}(\mathbf{E}_0 - \mathbf{E}))^k\mathbf{E}_0^{-1}$ and that each term in the sum is symmetric and positive semi-definite for $k > 0$ and positive definite for $k = 0$, it follows that $\mathbf{P}_1^l$ is symmetric positive definite. Using (4.43), we deduce that the sum converges as a Neumann series to $\mathbf{E}^{-1}$. Hence, it follows that

$$\mathbf{x}^\intercal\mathbf{E}_0^{-1}\mathbf{x} \leq \mathbf{x}^\intercal\mathbf{P}_1^l\mathbf{x} \leq \mathbf{x}^\intercal\mathbf{E}^{-1}\mathbf{x} \tag{4.46}$$

for all $\mathbf{x} \in \mathbb{R}^{n_S^+ n_R^+}$, which implies that $\mathbf{x}^\intercal\mathbf{E}\mathbf{x} \leq \mathbf{x}^\intercal(\mathbf{P}_1^l)^{-1}\mathbf{x} \leq \mathbf{x}^\intercal\mathbf{E}_0\mathbf{x}$ and, in turn,

$$(1-c)\mathbf{x}^\intercal(\mathbf{P}_1^l)^{-1}\mathbf{x} \leq \mathbf{x}^\intercal(\mathbf{E}-\mathbf{K})\mathbf{x} \leq \mathbf{x}^\intercal(\mathbf{P}_1^l)^{-1}\mathbf{x}, \tag{4.47}$$

where we used Lemma 4.3.3. The assertion follows then as in Section 4.3. □

Remark 4.5.5. *On the one hand, inspecting* (4.47) *we observe that the condition number of $\mathbf{P}_1^l(\mathbf{E}-\mathbf{K})$, and, similarly, of $\mathbf{E}^{-1}(\mathbf{E}-\mathbf{K})$ is $(1-c)^{-1}$, which is not robust for scattering dominated regimes $c \to 1$; cf. also Lemma 4.3.3. On the other hand, combining the second inequality in* (4.46) *with* (4.45)*, we obtain as in Lemma 4.2.1, that*

$$(1-(cg)^l)\mathbf{x}^\intercal(\mathbf{P}_1^l)^{-1}\mathbf{x} \leq \mathbf{x}^\intercal\mathbf{E}\mathbf{x} \leq \mathbf{x}^\intercal(\mathbf{P}_1^l)^{-1}\mathbf{x},$$

which shows that the condition number of $\mathbf{P}_1^l\mathbf{E}$ is bounded by $(1-(cg)^l)^{-1}$, which, for fixed $g < 1$, is robust for $c \to 1$.

Implementation of the subspace correction

The optimal subspaces for the correction (4.18) are constructed from the eigenfunctions associated with the largest eigenvalues of the generalized eigenproblem (4.20) as can be seen from the proof of Lemma 4.3.2 . The iterative computation of these eigenfunctions is, however, computationally expensive. Instead, we employ a different, computationally efficient tensor product construction that employs discrete counterparts of low-order spherical harmonics expansions motivated in Section 4.3. More precisely, the subspace for the correction is defined as $\mathbb{W}_{h,N}^+ = \mathbb{P}_{0,N}(\mathcal{T}_h^S) \otimes \mathbb{P}_1^c(\mathcal{T}_h^R)$, where $\mathbb{P}_{0,N}(\mathcal{T}_h^S) \subset \mathbb{P}_0(\mathcal{T}_h^S)$ is the space spanned by the eigenfunctions associated to the $d_N = (N+1)(N+2)/2$ largest eigenvalues of the generalized eigenvalue problem

$$\mathsf{S}^+\mathsf{W}^+ = \mathsf{M}^+\mathsf{W}^+\mathsf{\Lambda}^+$$

for the scattering operator, mimicking (4.7) after discretization. Note that d_N with N even is the number of even spherical harmonics of order less than or equal to N, and $\mathbb{P}_{0,N}(\mathcal{T}_h^S)$ approximates their span. Denote W_N^+ the corresponding matrix of coefficient vectors. The subspace $\mathbb{W}_{h,N}^+$ is spanned by the columns of the matrix $\mathbf{W}^+ = \mathsf{W}_N^+ \otimes \mathfrak{I}^+$. At the discrete level, the correction equation (4.18), thus, reads as

$$\big(\mathbf{W}^{+\intercal}(\mathbf{E}-\mathbf{K}^+)\mathbf{W}^+\big)\mathbf{u}_c = \mathbf{W}^{+\intercal}((\mathbf{E}-\mathbf{K}^+)\mathbf{P}_1 - \mathbf{I})((\mathbf{E}-\mathbf{K}^+)\mathbf{u}_n - \mathbf{q}). \tag{4.48}$$

The efficient assembly of the matrix on the left-hand side relies on the tensor product structure of $\mathbf{K}^+$ and the choice of $\mathbf{W}_N^+$ as outlined in the following. A simple and direct representation of the scattering operator on $\mathbb{W}_{h,N}^+$ is obtained by

$$\mathbf{W}^{+\intercal}\mathbf{K}^+\mathbf{W}^+ = \mathbf{\Lambda}_N^+ \otimes \mathfrak{M}_s^+.$$

Similarly, we have that $\mathbf{W}^{+\intercal}\mathbf{M}^+\mathbf{W}^+ = \mathbf{I}^+ \otimes \mathfrak{M}_t^+$, and the block-diagonal structure of $\mathbf{R}$ allows to compute $\mathbf{W}^{+\intercal}\mathbf{R}\mathbf{W}^+$, i.e. the (i,j)th block-entry is given by

$$\sum_{k=1}^{n_S^+} \mathfrak{R}_k(\mathbf{W}_N^+(k,i)\mathbf{W}_N^+(k,j))$$

which requires $O(n_S^+(n_R^+)^{(d-1)/d}d_N)$ many multiplications. The efficient assembly of the remaining term $\mathbf{W}^{+\intercal}\mathbf{A}^\intercal(\mathbf{M}^- - \mathbf{K}^-)^{-1}\mathbf{A}\mathbf{W}^+$ relies on another eigenvalue decomposition, which diagonalizes $\mathbf{M}^- - \mathbf{K}^-$ on the column range of $\mathbf{A}\mathbf{W}^+$. The arguments are similar to those in Section 4.5 and we leave the details to the reader.

4.6 Full algorithm and complexity

For the convenience of the reader we provide here the full algorithm of our numerical scheme. To simplify presentation we start with the application of $\mathbf{E}$ as given in Algorithm 1 and the application of $\mathbf{P}_1$ as given in Algorithm 2. The full preconditioned Richardson iteration (4.33) is outlined in Algorithm 3.

Algorithm 1 Apply $\mathbf{E}$, given a factorization of $\mathbf{S}_N^-$ as in (4.38).

function $\mathbf{y}$ =ApplyE($\mathbf{x}$)
  Solve $(\mathbf{M}^- - \mathbf{K}^-)\mathbf{z} = \mathbf{A}\mathbf{x}$ with PCG, preconditioned by $(\mathbf{M}^- - \mathbf{K}_N^-)^{-1}$ as in (4.39)
  $\mathbf{y} = \mathbf{A}^\intercal\mathbf{z} + \mathbf{M}^+\mathbf{x} + \mathbf{R}\mathbf{x}$
end function

Algorithm 2 Apply $\mathbf{P}_1 = \mathbf{P}_1^l$ as given in (4.44).

function $\mathbf{z}$ =ApplyP$_1$($\mathbf{x}$)
  $\mathbf{z} = 0$
  for $k = 0, 1, \ldots, l$ **do**
    $\mathbf{z} = \mathbf{z} - \mathbf{E}_0^{-1}$(ApplyE($\mathbf{z}$)$-\mathbf{x}$)
  end for
end function

For the efficient implementation of these algorithms one may exploit that, except for $\mathbf{R}$, all matrices provide a tensor product structure, see (4.29)–(4.31), allowing for efficient storage in $\mathcal{O}(n_S^\pm + n_R^\pm)$ or $\mathcal{O}(c_\mathcal{H} n_S^\pm + n_R^\pm)$ complexity by using their sparsity or their $\mathcal{H}^2$-matrix representation[1]. Here, $c_\mathcal{H}$ is a constant

[1]The storage requirements of $\mathbf{K}^+$ and $\mathbf{K}^-$ are $\mathcal{O}(c_\mathcal{H} n_S^\pm \log(n_S^\pm) + n_R^\pm)$ if $\mathcal{H}$-matrices are used instead of $\mathcal{H}^2$-matrices. In practice, $c_\mathcal{H}$ may depend on additional implementation dependent parameters, see [41, 40], which we neglect here for sake of simplicity.

Algorithm 3 Solve $\mathbf{E}\mathbf{u}^+ = \mathbf{K}^+\mathbf{u}^+ + \mathbf{q}$ according to (4.33)

Compute $\mathsf{S}_N^+ = \mathsf{M}^+\mathsf{W}_N^+\mathsf{\Lambda}_N^+(\mathsf{W}_N^+)^\intercal\mathsf{M}^+$
Compute $\mathsf{S}_N^- = \mathsf{M}^-\mathsf{W}_N^-\mathsf{\Lambda}_N^-(\mathsf{W}_N^-)^\intercal\mathsf{M}^-$

Compute $\mathbf{E}_c = \mathbf{W}^{+\intercal}(\mathbf{E} - \mathbf{K}^+)\mathbf{W}^+$ as in Section 4.5

Choose $\mathbf{u}_0^+$
for $n = 0, 1, 2, \ldots$ **do**

  $\mathbf{r} = \textsc{Apply}\mathbf{E}(\mathbf{u}_n^+) - \mathbf{K}^+\mathbf{u}_n^+ - \mathbf{q}$
  $\mathbf{s} = \textsc{Apply}\mathbf{P}_1(\mathbf{r})$
  $\mathbf{u}_{n+1/2}^+ = \mathbf{u}_n^+ - \mathbf{s}$ ▷ Half-step

  $\mathbf{q}_c = \mathbf{W}^{+\intercal}\big(\textsc{Apply}\mathbf{E}(\mathbf{s}) - \mathbf{K}^+\mathbf{s} - \mathbf{q} - \mathbf{r}\big)$
  Solve $\mathbf{E}_c\mathbf{u}_{n+1/2,c}^+ = \mathbf{q}_c$
  $\mathbf{u}_{n+1}^+ = \mathbf{u}_{n+1/2}^+ + \mathbf{W}^+\mathbf{u}_{n+1/2,c}^+$ ▷ Subspace correction
end for

related to the compression pattern of the $\mathcal{H}^2$-matrix. The storage requirements and application of $\mathbf{R}$ have complexity $\mathcal{O}(n_S^+(n_R^+)^{(d-1)/d})$. The relation (4.37) then allows for an efficient application of all matrices occurring in (4.28) in $\mathcal{O}(n_S^\pm n_R^\pm)$ or $\mathcal{O}(c_\mathcal{H} n_S^\pm n_R^\pm)$ operations. Since the solution vector itself has size $n_S^+ n_R^+$, see also (4.25), and since $3n_S^+ = n_S^-$ and $n_R^+ \sim n_R^-$, all matrices appearing in (4.28) can be stored and applied with linear complexity.

In the following we elaborate the algorithmic complexities of Algorithms 1 to 3 in more detail.

Complexity of applying E

The listing of Algorithm 1 directly indicates that the main effort of applying $\mathbf{E}$ lies in the preconditioned conjugate gradient method for applying $(\mathbf{M}^- - \mathbf{K}^-)^{-1}$. From Lemma 4.5.1, we obtain that $(\mathbf{M}^- - \mathbf{K}_N^-)^{-1}(\mathbf{M}^- - \mathbf{K}^-)$ is applicable in $\mathcal{O}((d_N + c_\mathcal{H})n_S^- n_R^-)$ operations, while its condition number is $(1-(cg)^{N+2})^{-1}$. This implies an iteration count for the application of $(\mathbf{M}^- - \mathbf{K}^-)^{-1}$ proportional to $(1-(cg)^{N+2})^{-1/2}$ for $cg \approx 1$ when using the preconditioned conjugate gradient method with a fixed tolerance. The overall complexity for applying $(\mathbf{M}^- - \mathbf{K}^-)^{-1}$ and, thus, also $\mathbf{E}$ is then $\mathcal{O}((d_N + c_\mathcal{H})n_S^- n_R^-/(1-(cg)^{N+2})^{1/2})$. We note that typically $d_N \ll c_\mathcal{H}$ for moderate N.

Complexity of applying the preconditioner $\mathbf{P}_1^l$

$\mathbf{P}_1^l$ consists of $l-1$ applications of $\mathbf{E}$ and l applications of $\mathbf{E}_0^{-1}$. Since $\mathbf{E}_0$ is block-diagonal with n_S^+ sparse blocks of size $n_R^+ \times n_R^+$, the application of $\mathbf{E}_0^{-1}$ can be performed in $\mathcal{O}(n_S^+(n_R^+)^\gamma)$ if the inversion of each block has $\mathcal{O}((n_R^+)^\gamma)$ complexity.

This amounts to $\mathcal{O}(l(d_N + c_\mathcal{H})n_S^+ n_R^+/(1-(cg)^{N+2})^{1/2} + l n_S^+(n_R^+)^\gamma)$ complexity for the application of $\mathbf{P}_1^l$. For moderate N, the subspace correction amounts to solving an elliptic system that is reminiscent of an order N spheri-

cal harmonics approximation, which can be solved efficiently with a conjugate gradient method preconditioned by a V-cycle geometric multigrid with Gauss-Seidel smoother, cf. [44].

Let us also remark that each diagonal block of $\mathbf{E}_0$ discretizes an anisotropic diffusion problem with a diffusion tensor $\sigma_t^{-1} \int_{K_S} s \cdot s^\intercal ds$ for $K_S \in \mathcal{T}_h^S$. The results reported in [20] indicate that such problems can be treated efficiently by multigrid methods with line smoothing allowing for $\gamma = 1$. A full analysis in the present context is out of the scope of this paper, but any method that gives $\gamma = 1$ allows to perform one step in the Richardson iteration (4.33) in linear complexity in the dimension of the solution vector. Although $\gamma > 1$, sparse direct solvers may work well, too, cf. Table 4.9.

Complexity of the overall iteration

We start our considerations by remarking that the truncated eigendecompositions of the smaller matrices $\mathbf{S}^+$ and $\mathbf{S}^-$ can be obtained by a few iterations of an iterative eigensolver. Once this is achieved, the computation of the reduced matrix $\mathbf{E}_c$ can be achieved in $O(n_S^+ n_R^+ d_N)$ operations, see Section 4.5. Thus, the offline cost for the construction of the preconditioners are $O(n_S^+ n_R^+ d_N)$. The discussion on the application of $\mathbf{E}$ and $\mathbf{P}_1$ shows that a single iteration of Algorithm 3 can be accomplished in $\mathcal{O}(l(d_N + c_{\mathcal{H}}) n_S^+ n_R^+ / (1-(cg)^{N+2})^{1/2} + l n_S^+ (n_R^+)^\gamma)$ operations.

Let us remark that in the case $\gamma = 1$ each iteration has linear complexity and it can be implemented such that it offers a perfect parallel weak scaling in $n_S^+ n_R^+$ as long as the number of processors is bounded by n_S^+ and n_R^+. To see this, we note that, with $\mathbf{R}$ being the only exception, we are only relying on matrix-vector products of matrices having tensor-product structure (or sums thereof). Using the identity (4.37), it is clear that these operations offer the promised weak scaling when these matrix-matrix products are accelerated by a parallelization over the rows and columns of the middle matrix. The matrix $\mathbf{R}$ does not directly provide such a structure, but its block diagonal structure, cf. (4.31), provides possibilities for a perfectly weakly scaling implementation as well.

In summary, each step in (4.33) can be executed very efficiently with straightforward parallelization. In the next section we show numerically that the number of iterations required to decrease the error below a given threshold is small already for small values of l and N.

4.7 Numerical realization and examples

We present the performance of the proposed iterative schemes using a lattice type problem [18], see fig. 4.1. Here, $R = (0,7) \times (0,7)$, the inflow boundary source $f = 0$, and $c = \|\sigma_s/\sigma_t\|_\infty \approx 0.999$. The coarsest triangulation of the sphere consists of 128 element, i.e., $n_S^+ = 64$, and $n_R^+ = 3\,249$ vertices to discretize the spatial domain. Finer meshes are obtained by uniform refinement; the new grid points for $\mathcal{T}_h^S$ are projected to the sphere. To minimize consistency errors, we use higher-order integration rules for the spherical integrals. The timings are performed using an AMD dual EPYC 7742 with 128 cores and with 1024GB memory.

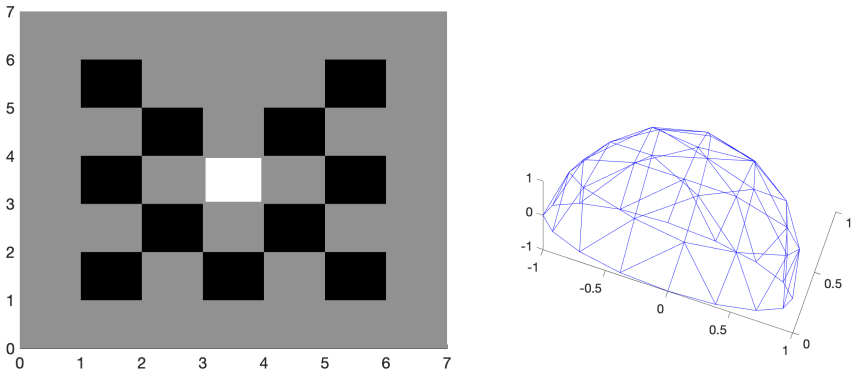


Figure 4.1: Left: geometry of the lattice problem. The optical parameters are $\sigma_s = 10$ and $\sigma_a = 0.01$ in the white and grey regions, $\sigma_s = 0$ and $\sigma_a = 1$ in the black regions and $q = 1$ in the grey region and $q = 0$ outside the grey region. Right: Sketch of the spherical grid.

Application of $(\mathcal{M} - \mathcal{K})^{-1}$

We show that $(\mathbf{M}^- - \mathbf{K}^-)^{-1}$ can be applied efficiently and robustly in g. To that end, we implemented a preconditioned conjugate gradient method with preconditioner $\mathbf{M}^- - \mathbf{K}_N^-$, see Section 4.5. Table Table 4.1 shows the required iteration counts to achieve a relative error below 10^{-13}. For all g, the iteration counts decrease with N as predicted by the considerations in Section 4.6. In particular, since $\mathbf{K}^- = \mathbf{K}_N^- = 0$, only one iteration is needed for convergence for $g = 0$. Moreover, we see that, although increasing the value of N increases the workload per iteration, the overall solution time can decrease, which is due to the fact that the scattering operator dominates the computational cost for moderate d_N, see Section 4.6. In the remainder of the paper, we employ $N = 5$, which yields fast convergence for the considered values of g.

Table 4.1: Iteration counts (timings in sec.) for the application of $(\mathbf{M}^- - \mathbf{K}^-)^{-1}$ using a preconditioned CG method with preconditioner $\mathbf{M}^- - \mathbf{K}_N^-$ and tolerance 10^{-13} for $n_S^+ = 256$ and $n_R^+ = 12\,769$.

		g					
N	d_N	0	0.1	0.3	0.5	0.7	0.9
-1	0	1 (1.6)	4 (4.2)	6 (6.1)	8 (8.0)	11 (10.7)	21 (19.9)
1	3	1 (1.6)	3 (3.3)	5 (4.9)	7 (6.7)	10 (9.5)	19 (17.4)
3	10	1 (1.7)	2 (2.6)	5 (5.0)	6 (6.1)	8 (7.8)	19 (17.6)
5	21	1 (1.8)	2 (2.7)	3 (3.6)	4 (4.5)	7 (7.1)	15 (14.2)
7	36	1 (1.8)	2 (2.8)	3 (3.6)	4 (4.6)	6 (6.4)	14 (13.4)
9	55	1 (1.9)	2 (2.8)	2 (2.8)	4 (4.7)	6 (6.4)	12 (12.1)

Convergence rates

We study the norm η of the iteration matrix $(\mathbf{I}^+ - \mathbf{P}_G)(\mathbf{I}^+ - \mathbf{P}_1^l(\mathbf{E} - \mathbf{K}^+))$ defined in (4.36) and its spectral radius

$$\rho = \max\{|\lambda| : \ \lambda \text{ is an eigenvalue of } (\mathbf{I}^+ - \mathbf{P}_G)(\mathbf{I}^+ - \mathbf{P}_1^l(\mathbf{E} - \mathbf{K}^+))\}$$

for different choices of preconditioners $\mathbf{P}_1 = \mathbf{P}_1^l$, anisotropy factors g and dimensions d_N chosen for the subspace correction. Since $\mathbf{P}_G$ is a projection, we have that

$$(\mathbf{I}^+ - \mathbf{P}_G)^\intercal(\mathbf{E} - \mathbf{K}^+)(\mathbf{I}^+ - \mathbf{P}_G) = (\mathbf{E} - \mathbf{K}^+)(\mathbf{I}^+ - \mathbf{P}_G).$$

Therefore, η^2 is the largest eigenvalue of the eigenvalue problem

$$(\mathbf{I}^+ - \mathbf{P}_1^l(\mathbf{E} - \mathbf{K}^+))(\mathbf{I}^+ - \mathbf{P}_G)(\mathbf{I}^+ - \mathbf{P}_1^l(\mathbf{E} - \mathbf{K}^+))\mathbf{w} = \lambda\mathbf{w}.$$

We use MATLAB's eigs function to compute ρ and η with tolerance 10^{-7} and maximum iterations set to 300.

For the isotropic case $g = 0$, $\mathbf{P}_1^l = \mathbf{E}_0^{-1} = \mathbf{E}^{-1}$, i.e., ρ and η do not depend on l. For $N = 0$, Table 4.2 shows that the values of η and ρ are essentially independent of the discretization parameters, see also [9]. We observed numerically that choosing $N \in \{2, 4\}$ improves the values of ρ and η only slightly.

Table 4.2: Values of ρ and η of the iteration matrix for $g = 0$ and different angular grids.

n_S^+	16	64	256	1024	4096
η	0.385	0.429	0.445	0.450	0.451
ρ	0.212	0.261	0.280	0.286	0.288

In the next experiments, we vary g from 0.1 to 0.9 in steps of 0.2. Table 4.3–Table 4.7 display the corresponding values of ρ and η. For these anisotropic cases, the iteration count l for the preconditioner $\mathbf{P}_1^l$ as well as the number d_N, defined in Section 4.5, play an important role. For all combinations of d_N and l, we observe a convergent behavior with $\eta \leq c < 1$, which is in line with Lemma 4.5.4. The values of ρ and η decrease substantially with increasing d_N which is inline with the motivation of Section 4.3, while, for fixed d_N a saturation in l can be observed. For d_N sufficiently large, it seems that $\rho = \eta = g^l$, see, e.g. Table 4.6 for d_4 and $1 \leq l \leq 4$. We may conclude that we can achieve very good convergence rates for moderate values of d_N and l if combined appropriately.

$\mathcal{H}^2$-matrix approximation of $\mathcal{S}$

We demonstrate the $\mathcal{H}^2$-compressibility of the scattering operator $\mathcal{S}$. Since every $\mathcal{H}^2$-matrix can be represented as an $\mathcal{H}$-matrix, this also demonstrates the compressibility of $\mathcal{S}$ by means of $\mathcal{H}$-matrices. For the implementation we use a MEX interface to include the library H2LIB [32] into our MATLAB-implementation.

For the numerical experiments themselves, we choose $g = 0.5$ and the same quadrature formula in our MATLAB implementation and in our implementation

Table 4.3: Values of ρ and η for $g = 0.1$ and different values of d_N and l to realize $\mathbf{P}_1^l$.

	$d_0 = 1$		$d_2 = 6$		$d_4 = 15$	
l	ρ	η	ρ	η	ρ	η
1	0.298	0.432	0.156	0.247	0.117	0.161
2	0.264	0.429	0.101	0.237	0.048	0.141
3	0.261	0.429	0.097	0.237	0.043	0.141
4	0.261	0.429	0.097	0.237	0.042	0.141
5	0.261	0.429	0.097	0.237	0.042	0.141
6	0.261	0.429	0.097	0.237	0.042	0.141

Table 4.4: Values of ρ and η for $g = 0.3$ and different values of d_N and l to realize $\mathbf{P}_1^l$.

	$d_0 = 1$		$d_2 = 6$		$d_4 = 15$	
l	ρ	η	ρ	η	ρ	η
1	0.392	0.473	0.311	0.332	0.300	0.302
2	0.299	0.448	0.146	0.246	0.106	0.157
3	0.284	0.447	0.111	0.242	0.060	0.146
4	0.281	0.447	0.103	0.242	0.050	0.146
5	0.280	0.447	0.101	0.242	0.047	0.146
6	0.280	0.447	0.101	0.242	0.046	0.146

Table 4.5: Values of ρ and η for $g = 0.5$ and different values of d_N and l to realize $\mathbf{P}_1^l$.

	$d_0 = 1$		$d_2 = 6$		$d_4 = 15$	
l	ρ	η	ρ	η	ρ	η
1	0.522	0.553	0.499	0.499	0.499	0.499
2	0.386	0.489	0.265	0.301	0.250	0.255
3	0.361	0.482	0.174	0.260	0.136	0.175
4	0.358	0.480	0.147	0.254	0.089	0.159
5	0.357	0.480	0.140	0.253	0.070	0.156
6	0.357	0.480	0.137	0.253	0.062	0.156

within the H2Lib. The compression algorithm of H2Lib uses multivariate polynomial interpolation, requiring the extension of the Henyey-Greenstein kernel as in (4.42). The compression parameters are set to an admissibility parameter $\eta_{\mathcal{H}} = 1.4$, $p = 4$ interpolation points on a single interval and a minimal block size parameter $n_{\min} = 64$, see [41, 40]. We also tested an implementation without the need for an extension within the Bembel library [34] which yields similar results, but requires a finite element discretization on quadrilaterals, rather than triangles. In both cases, the differences between dense and compressed scattering matrix are below the discretization error.

Table 4.8 lists the memory requirements, setup time, and time for a single matrix-vector multiplication of $\mathbf{S}^+$ in dense and $\mathcal{H}^2$-compressed form. We can

Table 4.6: Values of ρ and η for $g = 0.7$ and different values of d_N and l to realize $\mathbf{P}_1^l$. The symbol $-$ indicates that MATLAB's eigs function has not converged to the desired tolerance.

	$d_0 = 1$		$d_2 = 6$		$d_4 = 15$	
l	ρ	η	ρ	η	ρ	η
1	—	0.699	0.699	0.699	0.699	0.699
2	0.537	0.582	0.489	0.489	0.489	0.489
3	0.515	0.567	0.349	0.366	0.342	0.342
4	0.512	0.565	0.270	0.319	0.241	0.253
5	0.511	0.564	0.248	0.309	0.178	0.212
6	0.511	0.564	0.239	0.306	0.142	0.195

Table 4.7: Values of ρ and η for $g = 0.9$ and different values of d_N and l to realize $\mathbf{P}_1^l$. The symbol $-$ indicates that MATLAB's eigs function has not converged to the desired tolerance.

	$d_0 = 1$		$d_2 = 6$		$d_4 = 15$	
l	ρ	η	ρ	η	ρ	η
1	—	—	—	—	—	0.899
2	0.808	0.808	0.808	0.808	0.808	0.808
3	0.764	0.775	0.758	0.758	0.727	0.727
4	0.763	0.773	0.757	0.757	0.653	0.653
5	0.763	0.772	0.757	0.757	0.587	0.587
6	0.763	0.772	0.757	0.757	0.528	0.528

clearly observe the quadratic complexity for storage and matrix-vector multiplication of the dense matrices and the asymptotically linear complexity of the $\mathcal{H}^2$-matrices. The scaling of the assembly times for dense and $\mathcal{H}^2$-matrices seems to be worse than predicted by theory, which is possibly caused by memory issues. Nevertheless, the scaling of the $\mathcal{H}^2$-matrices for the assembly times is much better than the one for dense matrices.

Benchmark example

The viability of the preconditioned Richardson iteration (4.33) is shown for some larger computations. We fix $g = 0.5$ and solve the even-parity equations (4.32) for the lattice problem. We fix $l = 4$ steps to realize the preconditioner $\mathbf{P}_1^l$ and $N = 4$, i.e., we use $d_4 = 15$ eigenfunctions of $\mathbf{S}^+$ for the subspace correction, cf. Section 4.5. In view of Table 4.5, we expect a contraction rate $\eta \approx 0.16$. Therefore, in order to achieve an error bound $\|\mathbf{u}^+ - \mathbf{u}_n^+\|_{\mathbf{E}-\mathbf{K}^+} < 10^{-8}$, we expect to require $n \approx 10$ iterations. In our implementation, we choose $\mathbf{u}_0^+ = 0$, and we stop the iteration at index n for which

$$\|\mathbf{u}_n^+ - \mathbf{u}_{n-1}^+\|_{\mathbf{E}-\mathbf{K}^+} < 10^{-8}\|\mathbf{u}_1^+\|_{\mathbf{E}-\mathbf{K}^+}. \tag{4.49}$$

Note that, assuming a contraction rate $\eta = 0.16$, Banach's fixed point theorem asserts that the error satisfies $\|\mathbf{u}^+ - \mathbf{u}_n^+\|_{\mathbf{E}-\mathbf{K}^+} \leq 0.2\|\mathbf{u}_n^+ - \mathbf{u}_{n-1}^+\|_{\mathbf{E}-\mathbf{K}^+}$. The dimension of the problem on the finest grid is $n_R^+ n_S^+ = 207\,360\,000$, i.e., storing

Table 4.8: Memory consumption in MB, timings in sec. for assembly and matrix-vector multiplication of $\mathbf{S}^+$ and corresponding $\mathcal{H}^2$-matrix approximation $\overline{\mathbf{S}^+}$ for $g = 0.5$. Numbers in brackets indicate the ratio to the previous refinement level.

n_S^+	mem $\mathbf{S}^+$	setup $\mathbf{S}^+$	apply $\mathbf{S}^+$
64	0.0312	0.171	$6.9 \cdot 10^{-5}$
256	0.5 (16.0)	0.203 (1.2)	$6.5 \cdot 10^{-5}$ (0.9)
1 024	8 (16.0)	0.438 (2.2)	0.000313 (4.8)
4 096	128 (16.0)	4.2 (9.6)	0.00517 (16.5)
16 384	$2.05 \cdot 10^3$ (16.0)	189 (45.0)	0.0805 (15.6)
65 536	$3.28 \cdot 10^4$ (16.0)	$1.09 \cdot 10^4$ (57.5)	2.67 (33.1)
262 144	—	—	—
1 048 576	—	—	—
n_S^+	mem $\overline{\mathbf{S}^+}$	setup $\overline{\mathbf{S}^+}$	apply $\overline{\mathbf{S}^+}$
64	0.0313	0.00109	0.00025
256	0.502 (16.0)	0.0116 (10.7)	0.000547 (2.2)
1 024	11.3 (22.5)	0.139 (11.9)	0.0086 (15.7)
4 096	89.2 (7.9)	0.902 (6.5)	0.0841 (9.8)
16 384	484 (5.4)	4.75 (5.3)	0.328 (3.9)
65 536	$2.27 \cdot 10^3$ (4.7)	24.6 (5.2)	1.46 (4.4)
262 144	$9.53 \cdot 10^3$ (4.2)	182 (7.4)	6.92 (4.7)
1 048 576	$3.82 \cdot 10^4$ (4.0)	$1.46 \cdot 10^3$ (8.0)	28.5 (4.1)

the solution vector requires 1.5GB of memory. Note that the corresponding dimension of the solution vector to the mixed system is about 1.5×10^9. Motivated by Table 4.8 we implement the scattering operators $\mathbf{S}^+$ and $\mathbf{S}^-$ using dense matrices in this example. The application of $\mathbf{E}_0^{-1}$ is implemented with MATLAB's sparse LU factorization, i.e., here, $\gamma \leq 1.5$ in the complexity estimates of Section 4.6.

Figure 4.2 shows exemplary the spherical average of the computed solution for $n_S^+ = 1024$ and $n_R^+ = 12\,769$. Table 4.9 displays the iteration counts and timings for different grid refinements. We observe mesh-independent convergence behavior of the iteration, which matches well the theoretical bound $n \approx 10$. Furthermore, the computation time scales like $(n_R^+)^{1.3}$ for fixed n_S^+.

If n_S^+ increases from 1024 to 4096, the superlinear growth in computation time can be explained by using dense matrices for $\mathbf{S}^+$ and $\mathbf{S}^-$, which, as shown in Table 4.8, can be remedied by using the compressed scattering operators.

Table 4.9: Iteration index n (timings in sec.) such that (4.49) holds for the benchmark example.

	n_R^+		
n_S^+	3 249	12 769	50 625
64	8 (50)	9 (236)	9 (1 470)
256	9 (114)	9 (499)	9 (2 476)
1 024	9 (300)	9 (1 107)	10 (6 580)
4 096	9 (1 017)	9 (4 983)	10 (34 029)

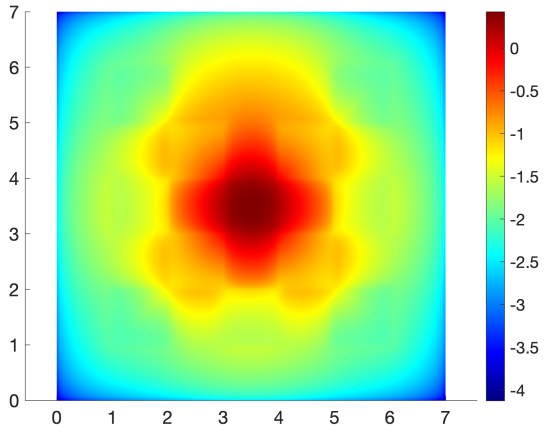


Figure 4.2: $\log_{10}$-plot of the spherical average of the numerical solution $\mathbf{u}^+$ to the benchmark problem as in Section 4.7 for $n_S^+ = 1\,024$ and $n_R^+ = 12\,769$.

4.8 Conclusions

We have presented efficient preconditioned Richardson iterations for anisotropic radiative transfer that are provably convergent and show robust convergence in the optical parameters, which comprises forwarded peaked scattering and heterogeneous absorption and scattering coefficients. This has been achieved by employing black-box matrix compression techniques to handle the scattering operator efficiently, and by construction of appropriate preconditioners. In particular, we have shown that, for anisotropic scattering, subspace corrections constructed from low-order spherical harmonics expansions considerably improve the convergence of our iteration.

On the discrete level, our preconditioners can be obtained algebraically from the matrices of any FEM code providing the matrices from the mixed system (4.28). We discussed further implementational details and their computational complexity, which, supported by several numerical tests, showed the efficiency of our method. If a solver with linear computational complexity for anisotropic elliptic problems is employed to realize $\mathbf{E}_0^{-1}$, each single iteration of our scheme has linear computational complexity in the discretization parameters. Our numerical examples employed low-order polynomials for discretization, but the presented methodology directly applies to high-order polynomial approximations as well.

Let us mention that the saddle-point problem (4.4) may also be solved using the MINRES algorithm after appropriate multiplication of the second equation by -1. In view of the inf-sup theory for (4.8)–(4.9) given in [7], block-diagonal preconditioners with blocks $\mathbf{E} - \mathbf{K}^+$ and $\mathbf{M}^- - \mathbf{K}^-$ lead to robust convergence behavior [24, Section 5.2], but the efficient inversion of $\mathbf{E} - \mathbf{K}^+$ is as difficult as solving the even-parity equations, which has been considered in this chapter.

Our subspace correction approach can also be related to multigrid schemes [4], and we refer to [25, 47, 16] and the references there in the context of radiative transfer. Comparing to non-symmetric Krylov space methods, such

as GMRES or BiCGStab, see [6, 50, 49] and the references there, our approach is very memory effective and monotone convergence behavior is guaranteed. Moreover, in view of its good convergence rates, the preconditioned Richardson iteration presented here is competitive to these multilevel and Krylov space methods.

Bibliography

[1] K. M. Case, P. F. Zweifel: Linear transport theory. Addison-Wesley, Reading (1967)

[2] E. E. Lewis, W. F. Miller Jr.: Computational methods of neutron transport. John Wiley & Sons, Inc., New York Chichester Brisbane Toronto Singapore (1984)

[3] J. Kópházi, D. Lathouwers: A space-angle DGFEM approach for the Boltzmann radiation transport equation with local angular refinement. Elsevier Journal of Computational Physics 297, 637–668 (2015)

[4] J. Xu, L. Zikatanov: The method of alternating projections and the method of subspace corrections in Hilbert space. Journal of the American Mathematical Society **15**(3), 573–597 (2002). doi:10.1090/s0894-0347-02-00398-3

[5] G. I. Marchuk, V. I. Lebedev: Numerical methods in the theory of neutron transport. Harwood Academic Publishers, Chur, London, Paris, New York (1986)

[6] M.L. Adams, E.W. Larsen: Fast iterative methods for discrete-ordinates particle transport calculations. Progress in Nuclear Energy 40(1), 3–159 (2002)

[7] H. Egger, M. Schlottbom: A mixed variational framework for the radiative transfer equation. Mathematical Models and Methods in Applied Sciences 22(03), [1150014] (2012)

[8] R. Becker, R. Koch, H.-J. Bauer, M.F. Modest: A finite element treatment of the angular dependency of the even-parity equation of radiative transfer. Journal of heat transfer 132(2) (2010)

[9] O. Palii, M. Schlottbom: On a convergent DSA preconditioned source iteration for a DGFEM method for radiative transfer. Elsevier, Computers & Mathematics with Applications 79(12), 3366–3377 (2020)

[10] J.S. Warsa, T.A. Wareing, J.E. Morel: Fully consistent diffusion synthetic acceleration of linear discontinuous S_N transport discretizations on unstructured tetrahedral meshes. Nuclear Science and Engineering **141**(3), 236–251 (2002). doi:10.13182/nse141-236

[11] T. A. Manteuffel, K. J. Ressel, G. Starke: A boundary functional for the least-squares finite-element solution for neutron transport problems. SIAM J.Numer. Anal. 2, 556–586 (2000)

[12] M.E. Tano, J.C. Ragusa: Sweep-net: An artificial neural network for radiation transport solves. Journal of Computational Physics **426**, 109757 (2021). doi:10.1016/j.jcp.2020.109757

[13] J.C. Ragusa, Y. Wang: A two-mesh adaptive mesh refinement technique for S_N neutral-particle transport using a higher-order DGFEM. Journal of Computational and Applied Mathematics **233**(12), 3178–3188 (2010). doi:10.1016/j.cam.2009.12.020

[14] Y. Wang, J.C. Ragusa: Diffusion synthetic acceleration for high-order discontinuous finite element S_N transport schemes and application to locally refined unstructured meshes. Nuclear Science and Engineering **166**(2), 145–166 (2010). doi:10.13182/nse09-46

[15] M.M. Crockatt, A.J. Christlieb, C.D. Hauck: Improvements to a class of hybrid methods for radiation transport: Nyström reconstruction and defect correction methods. Journal of Computational Physics **422**, 109765 (2020) doi:10.1016/j.jcp.2020.109765

[16] W. Shao, Q. Sheng, C. Wang: A cascadic multigrid asymptotic-preserving discrete ordinate discontinuous streamline diffusion method for radiative transfer equations with diffusive scalings. Computers & Mathematics with Applications **80**(6), 1650–1667 (2020). doi:10.1016/j.camwa.2020.08.002

[17] T.S. Haut, B.S. Southworth, P.G. Maginot, V.Z. Tomov: Diffusion synthetic acceleration preconditioning for discontinuous Galerkin discretizations of S_N transport on high-order curved meshes. SIAM Journal on Scientific Computing **42**(5), B1271–B1301 (2020). doi:10.1137/19M124993X

[18] T. A. Brunner: Forms of approximate radiation transport. In: Nuclear Mathematical and Computational Sciences: A Century in Review, A Century Anew Gatlinburg, American Nuclear Society, LaGrange Park, IL (2003), Tennessee, April 6-11, 2003

[19] V. Heningburg, C.D. Hauck: A hybrid finite-volume, discontinuous Galerkin discretization for the radiative transport equation. Multiscale Modeling & Simulation **19**(1), 1–24 (2021). doi:10.1137/19M1304520

[20] P. Hemker: Multigrid methods for problems with a small parameter in the highest derivative. In: D. Griffiths (ed.) Numerical Analysis, *Lecture Notes in Mathematics*, vol. 1066, pp. 106–121. Springer, Berlin Heidelberg (1984)

[21] H. Egger, M. Schlottbom: A perfectly matched layer approach for P_N-approximations in radiative transfer. SIAM Journal on Numerical Analysis **57**(5), 2166–2188 (2019). doi:10.1137/18M1172521

[22] C. Wang, Q. Sheng, W. Han: A discrete-ordinate discontinuous-streamline diffusion method for the radiative transfer equation. Communications in Computational Physics **20**(5), 1443–1465 (2018). doi:10.4208/cicp.310715.290316a

[23] W. Han, J. Huang, J.A. Eichholz: Discrete-ordinate discontinuous Galerkin methods for solving the radiative transfer equation. SIAM Journal on Scientific Computing **32**(2), 477–497 (2010). doi:10.1137/090767340

[24] A.J. Wathen: Preconditioning. Acta Numerica **24**, 329–376 (2015). doi:10.1017/s0962492915000021

[25] G. Kanschat, J.C. Ragusa: A robust multigrid preconditioner for S_NDG approximation of monochromatic, isotropic radiation transport problems. SIAM J. Sci. Comput. **36**(5), A2326–A2345 (2014). doi:10.1137/13091600X

[26] P. González-Rodríguez, A.D. Kim: Light propagation in tissues with forward-peaked and large-angle scattering. Applied Optics **47**(14), 2599 (2008). doi:10.1364/ao.47.002599

[27] X. Meng, S. Wang, G. Tang, J. Li, C. Sun: Stochastic parameter estimation of heterogeneity from crosswell seismic data based on the Monte Carlo radiative transfer theory. J. of Geophys. and Eng. **14**, 621–632 (2017)

[28] B. Ahmedov, M. Grepl, M. Herty: Certified reduced-order methods for optimal treatment planning. Mathematical Models and Methods in Applied Sciences **26**(04), 699–727 (2016). doi:10.1142/S0218202516500159

[29] T. Tarvainen, A. Pulkkinen, B.T. Cox, S.R. Arridge: Utilising the radiative transfer equation in quantitative photoacoustic tomography. In: A.A. Oraevsky, L.V. Wang (eds.) Photons Plus Ultrasound: Imaging and Sensing 2017 (2017). doi:10.1117/12.2249310

[30] S.R. Arridge, J.C. Schotland: Optical tomography: forward and inverse problems. Inverse Problems **25**(12), 123010 (2009). doi:10.1088/0266-5611/25/12/123010

[31] W. Dahmen, F. Gruber, O. Mula: An adaptive nested source term iteration for radiative transfer equations. Mathematics of Computation **89**(324), 1605–1646 (2020). doi:10.1090/mcom/3505

[32] S. Börm, Scientific Computing Group of Kiel University: H2Lib. `https://www.h2lib.org`

[33] R. Kriemann: HLib Pro. https://www.hlibpro.com

[34] J. Dölz, H. Harbrecht, S. Kurz, M. Multerer, S. Schöps, F. Wolf: Bembel: The fast isogeometric boundary element C++ library for Laplace, Helmholtz, and electric wave equation. SoftwareX **11**, 100476 (2020). doi:10.1016/j.softx.2020.100476

[35] W. Fong, E. Darve: The black-box fast multipole method. Journal of Computational Physics **228**(23), 8712–8725 (2009).doi:10.1016/j.jcp.2009.08.031

[36] W. Dahmen, H. Harbrecht, R. Schneider: Compression techniques for boundary integral equations. Asymptotically Optimal Complexity Estimates. SIAM Journal on Numerical Analysis **43**(6), 2251–2271 (2006)

[37] J. Dölz, H. Harbrecht, M. Peters: An interpolation-based fast multipole method for higher-order boundary elements on parametric surfaces. International Journal for Numerical Methods in Engineering **108**(13), 1705–1728 (2016). doi:10.1002/nme.5274

[38] H. Harbrecht, M. Peters: Comparison of fast boundary element methods on parametric surfaces. Computer Methods in Applied Mechanics and Engineering **261–262**, 39–55 (2013)

[39] L. Greengard, V. Rokhlin: A fast algorithm for particle simulations. Journal of Computational Physics **73**(2), 325–348 (1987)

[40] W. Hackbusch: Hierarchical matrices: algorithms and analysis. Springer, Heidelberg (2015)

[41] S. Börm: Efficient numerical methods for non-local operators. *EMS Tracts in Mathematics*, vol. 14. European Mathematical Society (EMS), Zürich (2010)

[42] J.L. Guermond, G. Kanschat, J.C. Ragusa: Discontinuous Galerkin for the radiative transport equation. In: Recent developments in discontinuous Galerkin finite element methods for partial differential equations, *IMA Vol. Math. Appl.*, vol. 157, pp. 181–193. Springer, Cham (2014)

[43] M.F. Modest: Radiative heat transfer, second edn. Academic Press, Amsterdam (2003)

[44] S. Arridge, H. Egger, M. Schlottbom: Preconditioning of complex symmetric linear systems with applications in optical tomography. Appl. Numer. Math. **74**, 35–48 (2013)

[45] E.D. Aydin, C.R.R. de Oliveira, A.J.H. Goddard: A finite element-spherical harmonics radiation transport model for photon migration in turbid media. Journal of Quantitative Spectroscopy & Radiative Transfer **84**, 247–260 (2004)

[46] H. Egger, M. Schlottbom: Stationary radiative transfer with vanishing absorption. Math. Mod. Meth. Appl. Sci. **24**, 973–990 (2014)

[47] B. Lee: A multigrid framework for S_N discretizations of the Boltzmann transport equation. SIAM Journal on Scientific Computing **34**(4), A2018–A2047 (2012)

[48] K.F. Evans: The spherical harmonics discrete ordinate method for three-dimensional atmospheric radiative transfer. Journal of the Atmospheric Sciences **55**(3), 429–446 (1998)

[49] J.S. Warsa, T.A. Wareing, J.E. Morel: Krylov iterative methods applied to multidimensional S_N calculations in the presence of material discontinuities. Tech. rep., Los Alamos National Laboratory (2002)

[50] M. Badri, P. Jolivet, B. Rousseau, Y. Favennec: Preconditioned Krylov subspace methods for solving radiative transfer problems with scattering and reflection. Computers & Mathematics with Applications **77**(6), 1453–1465 (2019)

[51] Z. Sun, C.D. Hauck: Low-memory, discrete ordinates, discontinuous Galerkin methods for radiative transport. SIAM Journal on Scientific Computing **42**(4), B869–B893 (2020). doi:10.1137/19M1271956

Chapter 5

Phase-space Discontinuous Galerkin approximation for the Radiative Transfer Equation

5.1 Introduction

We consider the numerical solution of the radiative transfer equation in slab geometry, which has several applications such as atmospheric science [1], oceanography [5], pharmaceutical powders [4] or solid state lightning [2]. Let us refer to [3] for a recent introduction. In view of available well-posedness results [7], it is natural to assume that the total cross section σ_t, which is the sum of the scattering cross section σ_s and the absorption cross section σ_a, is strictly positive. In this situation, the radiative transfer equation is equivalent to the following second-order form of radiative transfer equation with Robin boundary conditions [6, 8],

$$-\partial_z(\frac{\mu^2}{\sigma_t}\partial_z u) + \sigma_t u = \sigma_s \int_0^1 u(\cdot,\mu')\,d\mu' + f \quad \text{in } \Omega, \tag{5.1}$$

$$u + \frac{\mu}{\sigma_t}\partial_n u = g \quad \text{on } \Gamma. \tag{5.2}$$

Here, $u(z,\mu)$ corresponds to the even part of the solution of the radiative transfer equation for $(z,\mu) \in \Omega = (0,L)\times(0,1)$. Furthermore, $\partial_n u(0,\mu) = -\partial_z u(0,\mu)$ and $\partial_n u(L,\mu) = \partial_z u(L,\mu)$ are the normal derivatives of u on the boundary of the slab, defined as $\Gamma = \Gamma_0 \cup \Gamma_L$, where $\Gamma_z = \{z\}\times(0,1)$. The functions f and g model volume and boundary sources, respectively.

Due to the product structure of Ω, it seems natural to use separate discretization techniques for the spatial variable z and the angular variable μ. This is for instance done in the spherical harmonics method, in which a truncated Legendre polynomial expansion is employed to discretize μ [11]. The resulting coupled system of Legendre moments, which still depend on z, is then discretized for instance by finite differences or finite elements [11]. Another class of approximations consists of discrete ordinate methods which perform a collocation in μ and the integral (5.1) is approximated by a quadrature rule [11]. The resulting system of transport equations is then discretized for instance by finite differences [11] or discontinuous Galerkin methods [12, 9], and also spatially adaptive schemes have been used [10].

The content of this chapter has been submitted to Applied Numerical Mathematics journal as joint work of O. Palii and M. Schlottbom.

A major drawback of the independent discretization of the two variables z and μ is that a local refinement in phase-space is not possible. Such local refinement is generally necessary to achieve optimal schemes. For instance, in slab geometry, the solution can be non-smooth in the two points $(z, \mu) = (0, 0)$ and $(z, \mu) = (L, 0)$, which are exactly the two points separating the inflow from the outflow boundary. Although certain tensor-product grids can resolve this geometric singularity for the slab geometry, such as double Legendre expansions [11], they fail to do so for generic multi-dimensional situations. Moreover, local singularities of the solution due to the optical parameters or the source terms can in general not be resolved with optimal complexity.

Phase-space discretizations have been used successfully for radiative transfer in several applications, see, e.g., [13, 14, 15, 16] for slab geometry, [17] for geometries with spherical symmetries, or [18, 19] for more general geometries. Let us also refer to [20] for a phase-space discontinuous Galerkin method for the nonlinear Boltzmann equation. A non-tensor product discretization that combines ideas of discrete ordinates to discretize the angular variable with a discontinuous Petrov-Galerkin method to discretize the spatial variable has been developed in [21].

In this work, we aim to develop a numerical method for (5.1)–(5.2) that allows for local mesh refinement in phase-space and that allows for a relatively simple analysis and implementation. To accomplish this, we base our discretization on a partition of Ω such that each element in that partition is the Cartesian product of two intervals. Local approximations are then constructed from products of polynomials defined on the respective intervals. Since such partitions generically contain hanging nodes, global approximations are generally discontinuous. Therefore, we employ a symmetric interior penalty discontinuous Galerkin formulation. Besides the proper treatment of traces, which requires the inclusion of a weight function in our case, the analysis of the overall scheme is along the standard steps for the analysis of discontinuous Galerkin methods [22]. As a result, we obtain a scheme that enjoys an abstract quasi-best approximation property in a mesh-dependent energy norm. Our choice of meshes also allows to explicitly estimate the constants in auxiliary tools, such as inverse estimates and discrete trace inequalities. As a result, we can give an explicit lower bound on the penalty parameter required for discrete stability. Our theoretical results about accuracy and stability of the method are confirmed by numerical examples. Moreover, we show that adaptively refined grids are able to construct approximations to non-smooth solutions in optimal complexity.

The outline of the rest of the manuscript is as follows. In Section 5.2 we introduce notation and collect technical tools, such as trace theorems. In Section 5.3 we derive and analyze the discontinuous Galerkin scheme. Section 5.4 presents numerical examples confirming the theoretical results of Section 5.3. Section 5.5 shows that our scheme works well with adaptively refined grids. The chapter closes with some conclusions in Section 5.6.

5.2 Preliminaries

We denote by $L^2(\Omega)$ the usual Hilbert space of square integrable functions and denote the corresponding inner product by

$$(u, v) = \int_\Omega u(z, \mu)v(z, \mu)\, d(z, \mu).$$

Furthermore, we introduce the Hilbert space

$$V = \{v \in L^2(\Omega) : \ \mu\partial_z v \in L^2(\Omega)\},$$

which consists of square integrable functions for which the weighted derivative is also square integrable; see [7, Section 2.2]. We endow the space V with the graph norm

$$\|v\|_V^2 = \|v\|_{L^2(\Omega)}^2 + \|\mu\partial_z v\|_{L^2(\Omega)}^2, \qquad v \in V.$$

To treat the boundary condition (5.2), let us introduce the following inner product

$$\langle u, v\rangle = \int_\Gamma uv\, \mu\, d\mu = \int_0^1 \big(u(L, \mu)v(L, \mu) + u(0, \mu)v(0, \mu)\big)\mu\, d\mu,$$

and the corresponding space $L^2(\Gamma; \mu)$ of all measurable functions v such that

$$\|v\|_{L^2(\Gamma;\mu)}^2 = \langle v, v\rangle < \infty.$$

According to [7, Theorem 2.8], functions in V have a trace on Γ and

$$\|v\|_{L^2(\Gamma;\mu)} \le \frac{2}{\sqrt{1 - \exp(-L)}}\|v\|_V. \tag{5.3}$$

For the analysis of the numerical scheme, we provide a slightly different trace lemma.

Lemma 5.2.1. *Let $K = (z^l, z^r) \times (\mu^b, \mu^t) \subset \Omega$ for $0 \le z^l < z^r \le L$ and $0 \le \mu^b < \mu^t \le 1$. Let $F = \{z_F\} \times (\mu^b, \mu^t)$ with $z_F \in \{z^l, z^r\}$ be a vertical face of K. Then, for every $v \in V$ it holds that*

$$\int_F |v|^2\mu\, d\mu \le \left(\frac{\mu^t}{z^r - z^l}\|v\|_{L^2(K)} + 2\|\mu\partial_z v\|_{L^2(K)}\right)\|v\|_{L^2(K)}.$$

Proof. Without loss of generality, we assume that $z^l = z_F = 0$ and $z^r = h_z$. From the fundamental theorem of calculus, we obtain that

$$w(0, \mu) = w(z, \mu) - \int_0^z \partial_z w(y, \mu)dy.$$

Multiplication by μ and integration over K yields the inequality

$$h_z \int_F |w|\mu d\mu \le \int_K |w|\mu d(z, \mu) + \int_K \int_0^z \mu|\partial_z w(y, \mu)|dyd(z, \mu).$$

Setting $w = v^2$ in the previous inequality, observing that $|\mu\partial_z w| \le 2|(\mu\partial_z v)v|$ and applying the Cauchy-Schwarz inequality shows that

$$\int_F |v|^2\mu d\mu \le \int_K |v|^2 \frac{\mu}{h_z} d(z, \mu) + 2\|\mu\partial_z v\|_{L^2(K)}\|v\|_{L^2(K)},$$

which concludes the proof. □

Weak formulation and solvability

Performing the usual integration-by-parts, see also [8], the weak formulation of (5.1)–(5.2) is as follows.

Find $u \in V$ such that

$$a^e(u,v) = (f,v) + \langle g, v\rangle \quad \forall v \in V, \tag{5.4}$$

with bilinear form $a^e : V \times V \to \mathbb{R}$,

$$a^e(u,v) = (\frac{1}{\sigma_t}\mu\partial_z u, \mu\partial_z v) + (\sigma_t u, v) - (\sigma_s P u, v) + \langle u, v\rangle. \tag{5.5}$$

Here, for ease of notation, we use the scattering operator $P : L^2(\Omega) \to L^2(\Omega)$,

$$(Pu)(z,\mu) = \int_0^1 u(z,\mu')\, d\mu'.$$

Under the assumptions $0 \le \sigma_s, \sigma_t \in L^\infty(0,L)$, $\sigma_t - \sigma_s \ge c > 0$, $f \in L^2(\Omega)$ and $g \in L^2(\Gamma;\mu)$, the weak solution $u \in V$ of (5.4) exists, by the Lax-Milgram lemma, cf., e.g., [22, Lemma 1.4]. For $\mu > 0$ fixed, problem (5.1)-(5.2) reduces to an elliptic problem for $u(\cdot,\mu)$ and smoothness of $z \mapsto u(z,\mu)$ is governed by the smoothness of the data and the coefficients [27]. Therefore, since $f \in L^2(\Omega)$, we have that the flux $\frac{\mu}{\sigma_t}\partial_z u \in V$. In particular, for a.e. fixed $\mu > 0$, the flux $\frac{\mu}{\sigma_t}\partial_z u$ is continuous as a function of $z \in (0,L)$. We denote by

$$V_* = \{u \in V : \frac{\mu}{\sigma_t}\partial_z u \in V\} \tag{5.6}$$

the space of regular solutions u.

5.3 Discontinuous Galerkin scheme

In the following we will derive the numerical scheme to approximation solutions to (5.4). After introducing a suitable partition of Ω using quadtree grids and corresponding broken polynomial spaces, we can essentially follow the standard procedure for elliptic problems, cf. [22]. One notable difference is that we need to incorporate the weight function μ on the faces.

Mesh and broken polynomial spaces

Discontinuous Galerkin methods can be formulated for rather general meshes. In order to simplify the presentation, and subsequently the implementation, we consider quadtree meshes [28] as follows. Let $\mathcal{T}$ be a partition of Ω such that

$$K = (z_K^l, z_K^r) \times (\mu_K^l, \mu_K^r) \quad \forall\, K \in \mathcal{T},$$

for illustration see Figure 5.1. We denote the local mesh size by $h_K = z_K^r - z_K^l$.

Next, let us introduce some standard notation. Denote $\mathbb{P}_k$ the space of polynomials of one real variable of degree $k \ge 0$, and let the broken polynomial space V_h be denoted by

$$V_h = \{v \in L^2(\Omega) :\, v_{|K} \in \mathbb{P}_{k+1} \otimes \mathbb{P}_k \,\forall K \in \mathcal{T}\}. \tag{5.7}$$

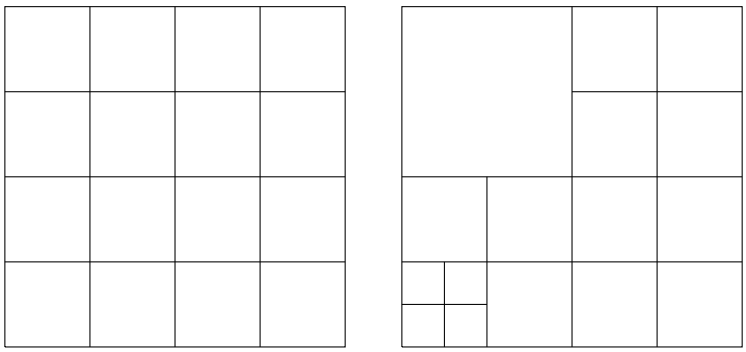


Figure 5.1: Left: Uniform mesh with 16 elements. Right: Non-uniform mesh with hanging nodes.

Moreover, let $V(h) = V + V_h$. By $\mathcal{F}_h^{vi}$ we denote the set of interior vertical faces, that is for any $F \in \mathcal{F}_h^{vi}$ there exist two disjoint elements

$$K_1 = (z_1^l, z_1^r) \times (\mu_1^l, \mu_1^r) \text{ and } K_2 = (z_2^l, z_2^r) \times (\mu_2^l, \mu_2^r)$$

such that $z_F = z_1^r = z_2^l$ and $F = \{z_F\} \times \big((\mu_1^l, \mu_1^r) \cap (\mu_2^l, \mu_2^r)\big)$. For $F \in \mathcal{F}_h^{vi}$ we define the jump and the average of $v \in V_h$ by

$$[\![v]\!] = v_{|K_1}(z_F, \mu) - v_{|K_2}(z_F, \mu), \qquad \{\!\{v\}\!\} = \frac{1}{2}(v_{|K_1}(z_F, \mu) + v_{|K_2}(z_F, \mu)).$$

Using the local mesh size h_{K_i}, $i \in \{1, 2\}$, of the element K_i in z-direction, we define the averaged mesh-size $H_F = (1/h_{K_1} + 1/h_{K_2})^{-1}$, which should not be confused with the length of the face F.

For an interior face $F \in \mathcal{F}_h^{vi}$ with $F = \{z_F\} \times (\mu_F^b, \mu_F^t)$, which is shared by two elements $K_F^i \in \mathcal{T}$, $i = 1, 2$, as above, let us introduce the sub-elements

$$E_F^i = (z_i^l, z_i^r) \times (\mu_F^b, \mu_F^t) \subset K_F^i. \tag{5.8}$$

We note that the inclusion in (5.8) can be strict in the case of hanging nodes, see for instance Figure 5.1.

Combining Lemma 5.2.1 with common inverse inequalities, cf. [29, Sect. 4.5], i.e., for any $k \geq 0$ there exists a constant $C_{ie}(k)$ such that

$$\left(\int_{z^l}^{z^r} |v'|^2 dz\right)^{1/2} \leq \frac{\sqrt{C_{ie}}}{z^r - z^l} \left(\int_{z^l}^{z^r} |v|^2 dz\right)^{1/2} \quad \forall v \in \mathbb{P}_k, \tag{5.9}$$

we obtain the following discrete trace lemma.

Lemma 5.3.1 (Discrete trace inequality)**.** *Let $K = (z_K^l, z_K^r) \times (\mu_K^l, \mu_K^r) \in \mathcal{T}$ and let $F = \{z_F\} \times (\mu_F^b, \mu_F^t) \in \mathcal{F}_h^v$ be such that $F \subset \partial K$. Then, for any $k \geq 0$ there holds*

$$\|v\|_{L^2(F;\mu)}^2 \leq \frac{C_{dt}}{h_K} \|v\|_{L^2((z_K^l, z_K^r) \times (\mu_F^b, \mu_F^t))}^2 \quad \forall v \in \mathbb{P}_k,$$

where $C_{dt}(k) = 1 + \sqrt{C_{ie}(k)}$, and C_{ie} is the constant in (5.9)*.*

Derivation of the DG scheme

In order to extend the bilinear form defined in (5.5) to the broken space V_h, we denote with ∂_z^h the broken derivative operator such that

$$(\frac{\mu^2}{\sigma_t}\partial_z^h u_h, \partial_z^h v_h) = \sum_{K\in\mathcal{T}} \int_K \frac{\mu^2}{\sigma_t}\partial_z u_h \partial_z v_h d(z,\mu)$$

for $u_h, v_h \in V_h$. In view of (5.4), let us then introduce the bilinear form

$$a_h^e(u,v) = (\frac{\mu^2}{\sigma_t}\partial_z^h u, \partial_z^h v) + (\sigma_t u, v) - (\sigma_s P u, v) + \langle u, v\rangle,$$

which is defined on $V(h)$. Note that a^e and a_h^e coincide on V. A routine calculation, cf. [22, Chapter 4], yields for any solution $u \in V_*$ to (5.1)–(5.2) and $v \in V_h$ that

$$a_h^e(u,v) = (f,v) + \langle g, v\rangle + \sum_{F\in\mathcal{F}_h^{vi}} \int_F \Big(\{\!\{ \frac{\mu}{\sigma_t}\partial_z^h u \}\!\} [\![v]\!] + [\![\frac{\mu}{\sigma_t}\partial_z^h u]\!] \{\!\{ v \}\!\} \Big) \mu \, d\mu.$$

Since $[\![\frac{\mu}{\sigma_t}\partial_z^h u]\!] = 0$ for all $F \in \mathcal{F}_h^{vi}$ by z-continuity of the flux of $u \in V_*$, we arrive at the identity

$$a_h^e(u,v) = (f,v) + \langle g, v\rangle + \sum_{F\in\mathcal{F}_h^{vi}} \int_F \{\!\{ \frac{\mu}{\sigma_t}\partial_z^h u \}\!\} [\![v]\!] \mu \, d\mu.$$

Hence, a consistent bilinear form is given by

$$a_h^c(u,v) = a_h^e(u,v) - \sum_{F\in\mathcal{F}_h^{vi}} \int_F \{\!\{ \frac{\mu}{\sigma_t}\partial_z^h u \}\!\} [\![v]\!] \mu \, d\mu,$$

which, for $V_{*h} = V_* + V_h$, is well-defined on $V_{*h} \times V_h$. Using that $[\![u]\!] = 0$ for any $u \in V$, we arrive at the following symmetric and consistent bilinear form

$$a_h^{cs}(u,v) = a_h^e(u,v) - \sum_{F\in\mathcal{F}_h^{vi}} \int_F \big(\{\!\{ \frac{\mu}{\sigma_t}\partial_z^h u \}\!\} [\![v]\!] + \{\!\{ \frac{\mu}{\sigma_t}\partial_z^h v \}\!\} [\![u]\!] \big) \mu \, d\mu,$$

which is again well-defined on $V_{*h} \times V_h$. We note that the summation over the vertical faces on the boundary Γ is included in the term $\langle u, v\rangle$ in a_h^e.

The stabilized bilinear form is then defined on $V_{*h} \times V_h$ by

$$a_h(u,v) = a_h^{cs}(u,v) + \sum_{F\in\mathcal{F}_h^{vi}} \frac{\alpha_F}{H_F} \int_F [\![u]\!] [\![v]\!] \mu d\mu,$$

with positive penalty parameter $\alpha_F > 0$, which will be specified below. Since $[\![u]\!] = 0$ on any $F \in \mathcal{F}_h^{vi}$ and $u \in V$, it follows that a_h is consistent, i.e., for $u \in V_*$ it holds

$$a_h(u, v_h) = a^e(u, v_h) \qquad \forall v \in V_h. \tag{5.10}$$

The discrete variational problem is formulated as follows:
Find $u_h \in V_h$ such that

$$a_h(u_h, v_h) = (f, v_h) + \langle g, v_h\rangle \quad \forall v_h \in V_h. \tag{5.11}$$

Analysis

For the analysis of (5.11), let us introduce mesh-dependent norms

$$\|v\|_{V_h}^2 = a_h^e(v,v) + \sum_{F\in\mathcal{F}_h^{v\,i}} H_F^{-1}\|[\![v]\!]\|_{L^2(F;\mu)}^2, \quad v\in V(h),$$

$$\|v\|_*^2 = \|v\|_{V_h}^2 + \sum_{F\in\mathcal{F}_h^{v\,i}} \frac{H_F}{C_{dt}}\|\{\!\{\frac{\mu}{\sigma_t}\partial_z^h v\}\!\}\|_{L^2(F;\mu)}^2, \quad v\in V_{*h}.$$

In order to show discrete stability and boundedness of a_h, we will use the following auxiliary lemma.

Lemma 5.3.2 (Auxiliary lemma)**.** *Let $F\in\mathcal{F}_h^{v\,i}$ be shared by the elements $K_F^1, K_F^2\in\mathcal{T}$. Then, for $w\in V_h$ and $v\in V(h)$ it holds that*

$$\int_F \{\!\{\frac{\mu}{\sigma_t}\partial_z^h w\}\!\}[\![v]\!]\mu d\mu \leq \frac{\sqrt{C_{dt}}}{2\sqrt{H_F}}\left(\|\frac{\mu}{\sigma_t}\partial_z w\|_{L^2(E_F^1)}^2 + \|\frac{\mu}{\sigma_t}\partial_z w\|_{L^2(E_F^2)}^2\right)^{1/2}\|[\![v]\!]\|_{L^2(F;\mu)},$$

with C_{dt} from Lemma 5.3.1 and sub-elements E_F^i, $i=1,2$, defined in (5.8)*.*

Proof. By definition of the average, we have that

$$\int_F \{\!\{\frac{\mu}{\sigma_t}\partial_z^h w\}\!\}[\![v]\!]\mu d\mu = \frac{1}{2}\int_F \frac{\mu}{\sigma_t}\partial_z w_1[\![v]\!]\mu d\mu + \frac{1}{2}\int_F \frac{\mu}{\sigma_t}\partial_z w_2[\![v]\!]\mu d\mu,$$

where w_1, w_2 denote the restrictions of w to K_F^1 and K_F^2, respectively. To estimate the first integral on the right-hand side, we employ the Cauchy-Schwarz inequality and Lemma 5.3.1 to obtain

$$\begin{aligned}\int_F \frac{\mu}{\sigma_t}\partial_z w_1[\![v]\!]\mu d\mu &\leq \|\frac{\mu}{\sigma_t}\partial_z w_1\|_{L^2(F;\mu)}\|[\![v]\!]\|_{L^2(F;\mu)}\\ &\leq \frac{\sqrt{C_{dt}}}{\sqrt{h_{K_F^1}}}\|\frac{\mu}{\sigma_t}\partial_z w_1\|_{L^2(E_F^1)}\|[\![v]\!]\|_{L^2(F;\mu)}.\end{aligned}$$

A similar estimate holds for the second integral. Hence, we can estimate

$$\int_F \{\!\{\frac{\mu}{\sigma_t}\partial_z w\}\!\}[\![v]\!]\mu d\mu \leq \frac{\sqrt{C_{dt}}}{2}\big(\|\frac{\mu}{\sigma_t}\partial_z w\|_{L^2(E_F^1)}^2 + \|\frac{\mu}{\sigma_t}\partial_z w\|_{L^2(E_F^2)}^2\big)^{1/2} \sqrt{\frac{1}{h_{K_F^1}}+\frac{1}{h_{K_F^2}}}\|[\![v]\!]\|_{L^2(F;\mu)},$$

which concludes the proof. □

The auxiliary lemma allows to bound the consistency terms in a_h, which gives discrete stability of a_h.

Lemma 5.3.3 (Discrete stability)**.** *For any $v\in V_h$ it holds that*

$$a_h(v,v) \geq \frac{1}{2}\|v\|_{V_h}^2$$

provided that $\alpha_F \geq 1/2 + C_{dt}$ with constant C_{dt} given in Lemma 5.3.1.

Proof. Let $v_h \in V_h$, and consider

$$a_h(v_h, v_h) = a_h^e(v_h, v_h) - 2 \sum_{F \in \mathcal{F}_h^{v\,i}} \int_F \{\!\{ \frac{\mu}{\sigma_t} \partial_z v_h \}\!\} [\![v_h]\!] \mu d\mu + \sum_{F \in \mathcal{F}_h^{v\,i}} \frac{\alpha_F}{H_F} \int_F [\![v_h]\!]^2 \mu d\mu.$$

Using Lemma 5.3.2, and the fact that each sub-element E_F^i touches at most two interior faces, an application of the Cauchy-Schwarz yields for any $\epsilon > 0$,

$$2 \sum_{F \in \mathcal{F}_h^{v\,i}} \int_F \{\!\{ \frac{\mu}{\sigma_t} \partial_z v_h \}\!\} [\![v_h]\!] \mu d\mu \leq \epsilon \| \frac{\mu}{\sigma_t} \partial_z^h v_h \|_{L^2(\Omega)}^2 + \sum_{F \in \mathcal{F}_h^{v\,i}} \frac{C_{dt}}{2\epsilon H_F} \int_F [\![v_h]\!]^2 \mu d\mu.$$

Hence, by choosing $\epsilon = 1/2$,

$$a_h(v_h, v_h) \geq \frac{1}{2} a_h^e(v_h, v_h) + \sum_{F \in \mathcal{F}_h^{v\,i}} \frac{\alpha_F - C_{dt}}{H_F} \int_F [\![v_h]\!]^2 \mu d\mu,$$

from which we obtain the assertion. □

Discrete stability implies that the scheme (5.11) is well-posed, cf. [22, Lemma 1.30].

Theorem 5.3.4 (Discrete well-posedness)**.** *Let $\alpha_F \geq 1/2 + C_{dt}$ with constant C_{dt} given in Lemma 5.3.1. Then for any $f \in L^2(\Omega)$ and $g \in L^2(\Gamma; \mu)$ there exists a unique solution $u_h \in V_h$ of the discrete variational problem* (5.11)*.*

Proof. The space V_h is finite-dimensional. Hence, Lemma 5.3.3 implies the assertion. □

To proceed with an abstract error estimate, we need the following boundedness result, which relies on the auxiliary Lemma 5.3.2.

Lemma 5.3.5 (Boundedness)**.** *For any $u \in V_{*h}$ and $v \in V_h$ it holds that*

$$a_h(u, v) \leq (C_{dt} + \alpha_F) \|u\|_* \|v\|_{V_h},$$

where α_F is as in Lemma 5.3.3.

Proof. We have that

$$\begin{aligned} a_h(u, v) = a_h^e(u, v) &- \sum_{F \in \mathcal{F}_h^{v\,i}} \int_F \{\!\{ \frac{\mu}{\sigma_t} \partial_z^h u \}\!\} [\![v]\!] \mu d\mu - \sum_{F \in \mathcal{F}_h^{v\,i}} \int_F \{\!\{ \frac{\mu}{\sigma_t} \partial_z^h v \}\!\} [\![u]\!] \mu \, d\mu \\ &+ \sum_{F \in \mathcal{F}_h^{v\,i}} \frac{\alpha_F}{H_F} \int_F [\![u]\!] [\![v]\!] \mu d\mu. \end{aligned}$$

The first two terms can be estimated using the Cauchy-Schwarz inequality as

$$a_h^e(u, v) \leq a_h^e(u, u)^{1/2} a_h^e(v, v)^{1/2},$$

$$\sum_{F \in \mathcal{F}_h^{v\,i}} \int_F \{\!\{ \frac{\mu}{\sigma_t} \partial_z^h u \}\!\} [\![v]\!] \mu d\mu \leq \sum_{F \in \mathcal{F}_h^{v\,i}} \| \{\!\{ \frac{\mu}{\sigma_t} \partial_z^h u \}\!\} \|_{L^2(F;\mu)} \| [\![v]\!] \|_{L^2(F;\mu)}.$$

For the third term, we use Lemma 5.3.2 to obtain

$$\sum_{F\in\mathcal{F}_h^{vi}} \int_F \{\!\!\{ \frac{\mu}{\sigma_t}\partial_z^h v \}\!\!\} [\![u]\!] \mu d\mu$$
$$\leq \sum_{F\in\mathcal{F}_h^{vi}} \frac{\sqrt{C_{dt}}}{2\sqrt{H_F}} \left(\|\frac{\mu}{\sigma_t}\partial_z v\|^2_{L^2(E_F^1)} + \|\frac{\mu}{\sigma_t}\partial_z v\|^2_{L^2(E_F^2)} \right)^{1/2} \|[\![u]\!]\|_{L^2(F;\mu)}.$$

To separate the terms that include u and v, respectively, we apply the Cauchy-Schwarz inequality once more and use again that each sub-element E_F^i touches at most two interior faces, to arrive at

$$a_h(u,v) \leq \Big(a_h^e(u,u) + \sum_{F\in\mathcal{F}_h^{vi}} \frac{H_F}{C_{dt}} \|\{\!\!\{ \frac{\mu}{\sigma_t}\partial_z^h u \}\!\!\}\|^2_{L^2(F;\mu)} + \frac{C_{dt}+\alpha_F}{H_F} \|[\![u]\!]\|^2_{L^2(F;\mu)} \Big)^{1/2}$$
$$\Big(a_h^e(v,v) + \frac{1}{2}\|\frac{\mu}{\sigma_t}\partial_z^h v\|^2_{L^2(\Omega)} + \sum_{F\in\mathcal{F}_h^{vi}} \frac{C_{dt}+\alpha_F}{H_F} \|[\![v]\!]\|^2_{L^2(F;\mu)} \Big)^{1/2},$$

which concludes the proof as $C_{dt} + \alpha_F \geq 3/2$. □

Combining, consistency, stability and boundedness ensures that the discrete solution u_h to (5.11) yields a quasi-best approximation to u, cf. [22, Theorem 1.35].

Theorem 5.3.6 (Error estimate)**.** *Let $f \in L^2(\Omega)$ and $g \in L^2(\Gamma;\mu)$, and denote $u \in V_*$ the solution to* (5.1)–(5.2) *and $u_h \in V_h$ the solution to* (5.11)*. Then the following error estimate holds true*

$$\|u - u_h\|_{V_h} \leq \big(1 + 2(C_{dt} + \alpha_F)\big) \inf_{v_h\in V_h} \|u - v_h\|_*,$$

provided that $\alpha_F \geq 1/2 + C_{dt}$.

5.4 Numerical examples

In the following we confirm the theoretical statements about stability and convergence of Section 5.3 numerically. Let $\sigma_s = 1/2$ and $\sigma_t = 1$ and the width of the slab be $L = 1$. We then define the source terms f and g in (5.1)–(5.2) such that the exact solution is given by the following function

$$u(z,\mu) = \big(1 + \exp(-\mu)\big)\chi_{\{\mu>1/2\}}(\mu)\exp(-z^2). \tag{5.12}$$

Here, $\chi_{\{\mu>1/2\}}(\mu)$ denotes the indicator function of the interval $(1/2, 1)$, i.e., u is discontinuous in $\mu = 1/2$, but note that $u \in V_*$. We compute the DG solution u_h of (5.11) on a sequence of uniformly refined meshes such that the initial mesh consists of 16 elements, see Figure 5.1. Hence, the discontinuity in u is resolved by the mesh. For our computations we use the lowest order space V_h with $k = 0$ in (5.7), that is piecewise constant functions in μ and piecewise linear functions in z. Using shifted Legendre polynomials, one can show that then $C_{ie} = 3$ in (5.9). In view of lemma 5.3.3, we choose $\alpha_F = 3/2 + \sqrt{3}$.

For the numerical solution of the resulting linear systems, we employ the usual source iteration [23]. Introducing the auxiliary bilinear form $b_h(u,v) =$

$a_h(u,v) - (\sigma_s P u, v)$, the source iteration performs the iteration $u_h^n \mapsto u_h^{n+1}$ by solving

$$b_h(u_h^{n+1}, v) = (\sigma_s P u_h^n, v) + (f, v) + \langle g, v\rangle \quad \forall v \in V_h. \tag{5.13}$$

The source iteration converges linearly with a rate σ_s/σ_t [23], which is bounded by $1/2$ in this example. For acceleration of the source iteration see also [23, 8]. The matrix representation of b_h has a block structure for the uniformly refined meshes considered in this example, and its inverse can be applied efficiently via LU factorization.

Table 5.1 shows the V_h-norm of the error $u - u_h$ between the exact and the numerical solution. As expected from of the polynomial degrees used for approximation, we observe linear convergence of the error in terms of the mesh size. For this example, we note that we found numerically a boundedness

Table 5.1: Error $\|u - u_h\|_{V_h}$ for uniformly refined mesh with N elements and solution u defined in (5.12).

N	16	64	256	1 014	4 096	16 384	65 536
$\|u - u_h\|_{V_h}$	0.0705	0.0352	0.0176	0.0088	0.0044	0.0022	0.0011

constant for the V_h-norm around $3.5 \leq C_{dt} + \alpha_F (\approx 4.23)$ and a coercivity constant larger than $0.75 \geq 1/2$, see Lemma 5.3.5 and Lemma 5.3.3.

5.5 Towards adaptive mesh refinement

In this section, we demonstrate that adaptive mesh refinement is beneficial if the non-smoothness of the solution is not resolved by the mesh. Different to the previous section, we assume for simplicity $\sigma_s = 0$ and

$$u(z,\mu) = \left(\mu^2 + \exp(-\mu)\chi_{\{\mu > 1/\sqrt{2}\}}(\mu)\right)\exp(-z^2). \tag{5.14}$$

The choice of $1/\sqrt{2}$ in the indicator function ensures that the corresponding discontinuity in u is never resolved exactly by our mesh. Note that again $u \in V_*$.

Figure 5.2 shows the convergence rate for uniformly refined meshes, adaptively refined meshes, and for comparison, the optimal rate $1/\sqrt{N}$ with N denoting the number of elements. We observe that the error for the uniformly refined grids behaves suboptimal, while the error for the adaptively refined grid is nearly parallel to the optimal curve. Here, we adapted the grid by using the local L^2-error between the numerical solution and the exact solution, i.e., for each $K \in \mathcal{T}$ we use

$$\eta_K^2 = \|u - u_h\|_{L^2(K)}^2,$$

which is computed using numerical quadrature; see Figure 5.2 for an illustration. The mesh is then refined by a Dörfler marking strategy [25], that is all elements in the set $\mathcal{K} \subset \mathcal{T}$ are refined, where $\mathcal{K} \subset \mathcal{T}$ is the set of smallest cardinality such that

$$\sum_{K\in\mathcal{K}} \eta_K^2 > 0.3 \sum_{K\in\mathcal{T}} \eta_K^2.$$

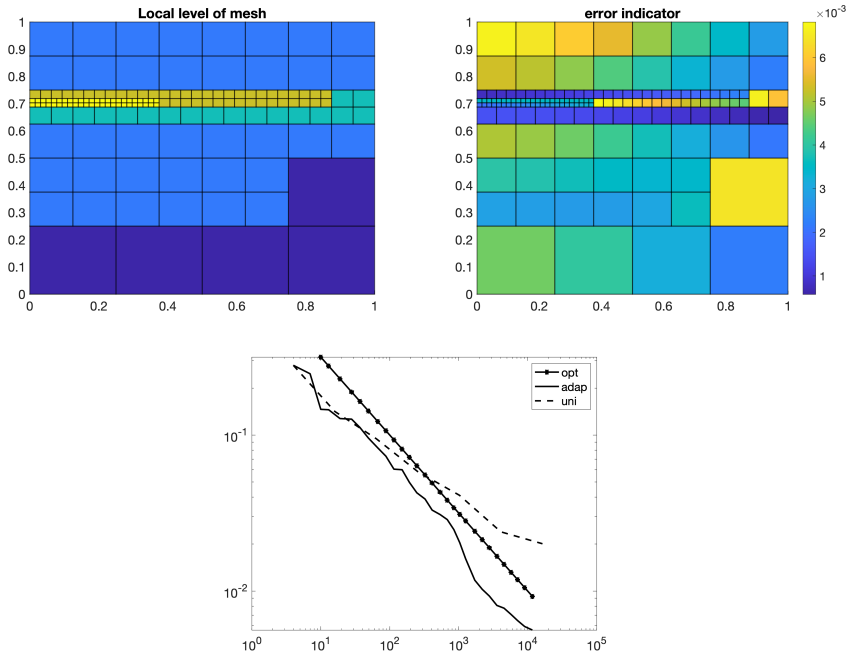


Figure 5.2: Non-smooth test case (5.14). Top left: Locally refined mesh with local mesh sizes varying from $1/2^2$ to $1/2^6$ for $N = 151$ elements. Top right: Local L^2-error times the size of an element for the grid shown left. Bottom: Convergence for uniformly refined grids (dotted), adaptively refined grids (connected), and, for comparison, the rate $1/\sqrt{N}$ (connected with stars) for different number of elements N in a double logarithmic scale.

An intermediate mesh with $N = 151$ elements obtained in this way is shown in Figure 5.2. We clearly see the local refinement towards the discontinuity of u for $\mu = 1/\sqrt{2}$.

5.6 Conclusions

We developed and analyzed a discontinuous Galerkin approximation for the radiative transfer equation in slab geometry. The use of quadtree-like grids allowed for a relatively simple analysis with similar arguments as for more standard elliptic problems. While such grids allow for local mesh refinement in phase-space, the implementation of the numerical scheme is straightforward. For sufficiently regular solutions, we showed optimal rates of convergence.

We showed by example that non-smooth solutions can be approximated well by adaptively refined grids. In order to automate the mesh adaptation procedure, an a posteriori error estimator is required. Since the solution to (5.1)–(5.2) is not in $H^1(\Omega)$ and is even allowed to be discontinuous, it seems difficult to generalize residual error estimators for elliptic problems, see, e.g., [22, Section 5.6] or [24, 25]. Upper bounds for the error can be derived for consistent approximations using duality theory [26]. Rigorous a posteriori error estimation has also been done using discontinuous Petrov-Galerkin discretiza-

tions [21]. We leave it to future research to investigate the construction of reliable and efficient local error estimators for the DG scheme considered here.

Bibliography

[1] J. E. Hansen, L. D. Travis: Light scattering in planetary atmospheres. Space science reviews, 16(4), 527-610 (1974)

[2] R. Melikov, D. A. Press, B. Ganesh Kumar, S. Sadeghi, S. Nizamoglu: Unravelling radiative energy transfer in solid-state lighting. Journal of Applied Physics, 123(2), 023103 (2018)

[3] R. Carminati, J. C. Schotland: Principles of Scattering and Transport of Light. Cambridge University Press (2021)

[4] T. Burger, J. Kuhn, R. Caps, J. Fricke: Quantitative determination of the scattering and absorption coefficients from diffuse reflectance and transmittance measurements: Application to pharmaceutical powders. Applied Spectroscopy, 51(3), 309-317 (1997)

[5] D. Arnush: Underwater light-beam propagation in the small-angle-scattering approximation. JOSA, 62(9), 1109-1111 (1972)

[6] H. Egger, M. Schlottbom: A mixed variational framework for the radiative transfer equation. Mathematical Models and Methods in Applied Sciences, 22(03), 1150014 (2012)

[7] V. Agoshkov, V.I. Agovskov: Boundary value problems for transport equations. Springer Science & Business Media (1998)

[8] O. Palii, M. Schlottbom: On a convergent DSA preconditioned source iteration for a DGFEM method for radiative transfer. Computers & Mathematics with Applications, 79(12), 3366-3377 (2020)

[9] J.-L. Guermond, G. Kanschat, J.C. Ragusa: Discontinuous Galerkin for the radiative transport equation. In Recent developments in Discontinuous Galerkin finite element methods for partial differential equations, Springer, 181–193 (2014)

[10] J.C. Ragusa, Y. Wang: A two-mesh adaptive mesh refinement technique for S_n neutral-particle transport using a higher-order DGFEM. Journal of computational and applied mathematics, 233(12), 3178–3188 (2010)

[11] J. J. Duderstadt, W. R. Martin: Transport theory. Transport theory (1979)

[12] W. Han, J. Huang, J. A. Eichholz: Discrete-ordinate discontinuous Galerkin methods for solving the radiative transfer equation. SIAM Journal on Scientific Computing, 32(2), 477-497 (2010)

[13] V.F. De Almeida: An iterative phase-space explicit discontinuous Galerkin method for stellar radiative transfer in extended atmospheres. Journal of Quantitative Spectroscopy and Radiative Transfer, 196, 254–269 (2017)

[14] L.H. Liu: Finite element solution of radiative transfer across a slab with variable spatial refractive index. International journal of heat and mass transfer, 48(11), 2260–2265 (2005)

[15] W.R. Martin, J.J. Duderstadt: Finite element solutions of the neutron transport equation with applications to strong heterogeneities. Nuclear Science and Engineering, 62(3), 371-390 (1977).

[16] W.R. Martin, C.E. Yehnert, L. Lorence, J.J. Duderstadt: Phase-space finite element methods applied to the first-order form of the transport equation. Annals of Nuclear Energy, 8(11-12), 633–646 (1981)

[17] D. Kitzmann, J. Bolte, A.B.C. Patzer: Discontinuous Galerkin finite element methods for radiative transfer in spherical symmetry. Astronomy & Astrophysics,595:A90 (2016)

[18] Y. Favennec, T. Mathew, M.A. Badri, P. Jolivet, B. Rousseau, D. Lemonnier, P.J. Coelho: Ad hoc angular discretization of the radiative transfer equation. Journal of Quantitative Spectroscopy and Radiative Transfer, 225, 301-318 (2019)

[19] J. Kópházi, D. Lathouwers: A space–angle DGFEM approach for the Boltzmann radiation transport equation with local angular refinement. Journal of Computational Physics, 297, 637–668 (2015)

[20] G. Kitzler, J. Schöberl: A high order space–momentum discontinuous Galerkin method for the Boltzmann equation. Computers & Mathematics with Applications, 70(7), 1539–1554 (2015)

[21] W. Dahmen, F. Gruber, O. Mula: An adaptive nested source term iteration for radiative transfer equations. Mathematics of Computation, 89(324), 1605–1646 (2020)

[22] D.A. Di Pietro, A. Ern: Mathematical aspects of discontinuous Galerkin methods. 69 Mathématiques & Applications (Berlin) [Mathematics & Applications]. Springer, Heidelberg (2012)

[23] M. L. Adams, E. W. Larsen: Fast iterative methods for discrete-ordinates particle transport calculations. Progress in nuclear energy, 40(1), 3-159 (2002)

[24] M. Ainsworth, J. T. Oden: A posteriori error estimation in finite element analysis. Computer methods in applied mechanics and engineering, 142(1-2), 1-88 (1997)

[25] R. Verfürth: A posteriori error estimation techniques for finite element methods. OUP Oxford (2013)

[26] W. Han: A posteriori error analysis in radiative transfer. Applicable Analysis, 94(12), 2517-2534 (2015)

[27] D. Gilbarg, N. S. Trudinger: Elliptic partial differential equations of second order. Classics in Mathematics. Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition.

[28] P. J. Frey, P. L. George: Mesh generation: application to finite elements. John Wiley & Sons Inc., 2nd edition (2008)

[29] S. C. Brenner, L. R. Scott: The mathematical theory of finite element methods (Vol. 3). Springer (2008)

Summary

In this thesis we studied approximation methods for the radiative transfer equation, which has numerous important applications, see Chapter 1. For most of these applications the radiative transfer equation cannot be solved analytically and a wide variety of numerical methods has been developed.

As an important introductory step, we gave an overview of classical semi-discretizations in the angular component. The two most frequently used discretizations - the discrete ordinates and spherical harmonics methods are summarised in Chapter 1. While the spherical harmonics discretization allows to turn the radiative transfer equation into a system of linear equations with tridiagonal structure, approximating the boundary conditions effectively requires extra steps. Furthermore, since the spherical harmonics expansion is a global approximation method, it is not suited for approximating non-smooth or discontinuous solutions, unlike the discrete ordinates method, which is a local approximation in the angle. The discrete ordinates method allows to obtain a consistent discretization of both the radiative transfer equation and the boundary conditions, though yielding dense scattering matrix.

A number of iterative techniques has been developed to tackle the difficulty that arises from the dense scattering matrix. A summary of some important methods was given in Chapter 2. We discussed two closely related methods - the first collision source method and the standard source iteration method, which is often accompanied by further preconditioning techniques, such as the diffusion synthetic acceleration technique. In the first collision source method the radiative transfer boundary value problem is split into two equations for the uncollided and collided components, which can be separately approximated by different numerical methods. In general this can, however, introduce consistency errors, which are difficult to analyse. We turned, therefore, to the source iteration method, which can be discretized consistently and for which convergence results are available. On the basis of these methods we gave in Chapter 2 a description of a splitting technique, which is essentially an extension of the first collision source method. These iterative methods, together with the aforementioned discretization techniques, provided an inspiration for the major part of our research.

In Chapter 3 we presented a discontinuous approximation in angle that allows for arbitrary partitions of the angular domain and arbitrary polynomial degrees on each element of that partition. As such, it can be understood as a generalization of the classical spherical harmonics approximation, where the angular domain is discretized by a single interval $(-1, 1)$ and polynomials of high degree, and the discrete ordinates method, where the angular domain is partitioned into several intervals with piecewise constant functions. In partic-

ular, the approach described in Chapter 3 allows to account for the natural discontinuity of the solution at $\mu = 0$. In Chapter 3 an *hp*-discretization was applied to the even-parity formulation of the radiative transfer equation with isotropic scattering. Moreover, we developed and analysed an iterative solution technique that employs subspace correction as a preconditioner. Our approach was inspired by the DSA preconditioned source iteration method. It was shown that our iterative method exhibited convergence independent of the resolution of the computational mesh.

In Chapter 4 we then focused on an efficient iterative framework that is capable of accurately solving the system of linear equations that arises from the discretization of anisotropic radiative transfer problems. In case of forward-peaked scattering the convergence of the standard DSA-preconditioned source iteration method is slow, hence acceleration with the use of an appropriate preconditioning technique is necessary. In Chapter 4 we proposed a provably convergent iterative method, equipped with two preconditioners, one of which corresponds to the efficient approximate inversion of transport. The second preconditioner was used to improve the standard contraction rate $\|\frac{\sigma_s}{\sigma_t}\|_\infty$ in the source iteration method. The subspace correction is then constructed from low order spherical harmonics expansions - eigenfunctions corresponding to the largest eigenvalues of the anisotropic scattering operator. The method is shown to be efficient if the scattering operator is applied properly, for which we used $\mathcal{H}$ and $\mathcal{H}^2$-matrix compression algorithms.

Finally, in Chapter 5 we considered a non-tensor product discontinuous Galerkin discretization for the even-parity radiative transfer equations for the slab geometry. We proved stability and well-posedness for the symmetric interior penalty discontinuous Galerkin method. We also investigated the numerical convergence of the phase-space discontinuous Galerkin method. For piecewise smooth solutions the phase-space discontinuous Galerkin method with low order polynomials displays a linear rate of convergence. We show numerically that in case of non-smooth solutions the use of adaptive mesh refinement allows for efficient approximation. The question of an appropriate choice of the error estimator remains an open question. Despite the similarity of the even-parity form of the RTE to standard elliptic problems, standard elliptic residual-based error estimators can not be generalized directly.

There are several problems, which are open for future research.

1. Developing and analyzing proper a-posteriori error estimators for our discontinuous Galerkin discretization of the radiative transfer equations in phase-space, allowing for *hp*-adaptivity.

2. Improving the error analysis of the preconditioned iterative schemes presented in Chapters 3 and 4, by proving precise quantitative rates of convergence.

3. Developing multigrid methods as an alternative to the preconditioned source iteration method.

Samenvatting

Dit proefschrift is gewijd aan numerieke methoden voor het oplossen van de stralingsoverdrachtvergelijking. Deze vergelijking heeft talloze toepassingen, zie Hoofdstuk 1. Voor de meeste toepassingen kan de stralingsoverdrachtvergelijking echter niet analytisch opgelost worden en daarom is een breed scala aan numerieke methoden ontwikkelt.

Als een eerste inleiding geven we een overzicht van een aantal semi-klassieke discretisaties in de hoekvariabele. De twee meest frequent gebruikte numerieke discretisaties: de "discrete ordinates" methode en de spectrale harmonische methode, die samengevat worden in Hoofdstuk 1, maken het mogelijk om de stralingsoverdrachtvergelijking te transformeren in een lineair stelsel van vergelijkingen met een drie-diagonale structuur. Het opleggen van de randcondities is echter niet triviaal. Aangezien de spectrale harmonische methode een globale benaderingstechniek is, is deze methode niet geschikt om niet-gladde of discontinue oplossingen te berekenen. Met de "discrete ordinates" methode is het well mogelijk om een consistente numerieke discretisatie van zowel de stralingsoverdrachtvergelijking als de randcondities te verkrijgen. Dit resulteert echter wel in een volle verstrooiingsmatrix.

Om de problemen op te lossen die veroorzaakt worden door de volle verstrooiingsmatrix zijn een aantal iteratieve numerieke methoden ontwikkeld. Een samenvatting van de belangrijkste methoden wordt gegeven in Hoofdstuk 2. Twee sterk gerelateerde methoden zijn de "first collision source" methode en de standaard "source iteration" methode, die vaak gebruikt wordt samen met preconditioneringstechnieken, zoals de "diffusion synthetic acceleration technique". In de "first collision source" methode wordt het randwaardeprobleem voor de stralingsoverdrachtvergelijking gesplitst in twee vergelijkingen, één voor de componenten die botsen en één voor de componenten die niet botsen, die ieder apart benaderd worden met een numerieke discretisatie. Dit resulteert echter vaak in consistentiefouten, die moeilijk zijn te analyseren. De voorkeur gaat daarom uit naar de "source iteration" methode, die op een consistentie manier numeriek benaderd kan worden en waarvoor convergentieresultaten beschikbaar zijn. Uitgaande van deze twee methoden geven we in Hoofdstuk 2 een beschrijving van een splitsingsmethode die in principe een uitbreiding is van de "first collision source" methode. Deze iteratieve methode, samen met de eerdergenoemde numerieke methoden, verschaffen de belangrijkste inspiratie voor een belangrijk deel van het onderzoek in dit proefschrift.

Hoofdstuk 3 is gewijd aan discontinue benaderingen in de stralingshoek. Hierbij is een willekeurige partitie van de stralingshoeken in verschillende domeinen toestaan en in ieder element van die partitie kunnen polynomen van verschillende graad worden gebruikt. Deze aanpak kan beschouwd worden als een

generalisatie van de klassieke bol-harmonische benaderingstechniek, waarin het hoekdomein wordt benaderd met het interval $(-1, 1)$ en polynomen van hoge graad, en de "discrete ordinates" methode waarin het hoekdomein verdeeld is in verschillende intervallen met stuksgewijs constante functies. Een belangrijk voordeel van de methode die besproken wordt in Hoofdstuk 3 is dat dit de mogelijkheid geeft om de natuurlijke discontinuïteit in de oplossing voor $\mu = 0$ goed te representeren. In Hoofdstuk 3 wordt een *hp*-discretisatie gebruikt voor de even-oneven formulering van de stralingsoverdrachtvergelijking met homogene verstrooiing. Daarnaast worden iteratieve methoden ontwikkeld en geanalyseerd die gebruik maken van deelruimte correcties als preconditioner. Onze aanpak was hierbij geïnspireerd door de DSA gepreconditioneerde "source iteration" methode. We laten zien dat de convergentie van onze iteratieve methode onafhankelijk is van de resolutie van het rekenrooster.

In Hoofdstuk 4 onderzoeken we een efficiënte klasse van iteratieve methoden die geschikt is om nauwkeurig het stelsel van lineaire vergelijkingen op te lossen dat voortkomt uit de numerieke discretisatie van de anisotrope stralingsoverdrachtvergelijking. De convergentie van de standaard DSA-gepreconditioneerde "source iteration" methode is in het geval van voorwaarts-gepiekte verstrooiing traag. Hierdoor is versnelling van de convergentie via het gebruik van preconditioneringsmethoden noodzakelijk. In Hoofdstuk 4 introduceren we een iteratieve methode waarvoor we de convergentie kunnen bewijzen. Deze methode bestaat uit twee preconditioners, waarvan één overeenkomt met een efficiënte benadering van de inverse van de transport termen in de vergelijking. De tweede preconditioner wordt gebruikt om de standard contractie coëfficiënt $\|\frac{\sigma_s}{\sigma_t}\|_\infty$ in de "source iteration" methode te verbeteren. De deelruimte correctie wordt geconstrueerd via een bol-harmonische reeks met een gering aantal termen, die overeenkomt met de eigenfuncties die gerelateerd zijn aan de grootste eigenwaarden van de anisotrope verstrooiingsoperator. Deze methode is efficiënt wanneer de verstrooiingsoperator correct wordt toegepast, wat gedaan wordt via $\mathcal{H}$ and $\mathcal{H}^2$-matrix compressie algoritmes.

In Hoofdstuk 5 bestuderen we discontinue Galerkin discretisaties voor de even-oneven partitie van de stralingsoverdrachtvergelijking in een plaatgeometrie. De basisfuncties zijn niet gebaseerd op een tensorproduct van basisfuncties. We bewijzen stabiliteit en goedgesteldheid voor de symmetrische "interior penalty" discontinue Galerkin methode. We hebben daarbij ook de numerieke convergentie van de "phase-space" discontinue Galerkin methode onderzocht. Voor stuksgewijs gladde oplossingen convergeert de "phase-space" discontinue Galerkin methode met lage orde polynomen lineair. We tonen via numerieke experimenten aan dat in het geval van niet-gladde oplossingen adaptieve methoden die gebaseerd zijn op lokale roosterverfijning effectief kunnen zijn. De juiste keuze van een foutenschatter is echter een open vraag. Ondanks de overeenkomst van de even-pariteit vorm van de stralingsoverdrachtvergelijking met standaard elliptisch problemen kan een residu gebaseerde foutenschatter niet direct gegeneraliseerd worden naar de stralingsoverdrachtvergelijking.

Verschillende problemen staan nog open voor toekomstig onderzoek.

1. De ontwikkeling en analyse van geschikte a-posteriori foutenschatters voor discontinue Galerkin discretisaties van de stralingsoverdrachtvergelijking in de toestandsruimte die toepasbaar zijn voor *hp* adaptatie.

2. Het verbeteren van de foutenanalyse van de gepreconditioneerde iteratieve methoden, die besproken worden in Hoofdstukken 3 en 4, door het verkrijgen van een nauwkeurige kwantitatieve schatting van de convergentiesnelheid.

3. Het ontwikkelen van multigrid methoden als alternatief voor de gepreconditioneerde "source iteration" methode.

Acknowledgements

Dear friends and colleagues, the past 4 years have been a wonderful experience full of personal challenges and discoveries. I was fortunate to have worked with my supervisors Jaap van der Vegt and Matthias Schlottbom and to have been a part of the Mathematics of Computational Science (MACS) group.

Throughout all 4 years of me working on this project I have received a lot of support and guidance with writing and presenting, for which I want to thank my supervisors. They dedicated their time for reading my dissertation thoroughly and always provided particularly useful recommendations to make it better.

Jaap has contributed a lot to this joint work by asking the right questions at every presentation I have given for the group as well as by suggesting possible solutions based on his immense experience.

Matthias has been my daily supervisor for the whole duration of the project, I want to thank him for being an understanding mentor and for closely monitoring my progress. It is hard to count the number of hours we have spent in discussions at the whiteboard, brainstorming the solutions to the problems together. He did not only teach me many topic related things, but paid a lot of attention to the methods of presenting the work in a way most suitable for the reader. In the last year of the project we often talked about my future professional development. Finding a career path is not always straightforward, so I found Matthias's input to be very valuable.

I want to thank my group members for the wonderful atmosphere in the group and a lot of useful discussions we have had over the years. When I started at MACS I was greeted by Lars, Nishant, Poorvi, Sjoerd G. and Sjoerd. Lars, thank you for being such a good friend to me, I could discuss anything with you. You have always been open to discussions and questions, your incredible ability to explain difficult concepts has helped me so much over the years. Some of the moments I most enjoyed were playing various board games when we were in Zeist or London, thank you for suggesting Saboteur. Thank you for agreeing to be my paranymph! Nishant, thank you for our conversations about everything in the world, there was no topic we wouldn't talk about over the lunch. You were always honest with me about your views. I remember our train conversation about the open systems (after the seminar), it was fun. Poorvi, thank you for being there for me in some of the most difficult times of my life. You were one of my very first friends in the Netherlands and I treasure you for being always honest and direct with me. My English was not all too good in the beginning but your influence helped me to get better at it. You were a wonderful housemate, sharing a living space with you helped me to learn to be more understanding and responsible. Sjoerd G., thank you

for being so professional and knowledgeable, your expertise often helped me to advance my own knowledge on the subject. Sjoerd, thank you for being a kind and responsible person, and for the nice lunch conversations.

I would also want to thank other young researchers from MACS, who we have managed to keep contact with despite the pandemic of 2019-2021. In particular I want to thank Elena, Fengna, Hamza, Kaifang, Marek, Riccardo, Vincent and Xiangyi. It has been wonderful working with you, you helped create a friendly and positive environment in the group.

Let me also note the contribution of other members of Systems, Analysis and Computational Sciences (SACS) group. I want to thank Bernard, Christoph, Hans, Hil, Jurgen, Kathrin, Len, Manu, Mir, Pranab, Sahar, Stephan, Tracy, Tugce and Yoeri. All of you have contributed to my development as a researcher and inspired me with your own successes. Sahar, thank you for your friendship, we have shared a lot of conversations about our projects and future careers. It has been a pleasure for me. Tugce, thank for being a kind and close friend, talking to you often helped me in the moments of doubt. Thank you for sharing your teaching experience with me as well.

I would like to thank every friend and colleague of mine whose name was not mentioned in these Acknowledgements.

Last but not least, I want to thank the dearest and the closest people to me, who supported me every step of the way - my family in Ukraine, Netherlands and Greece. My mom Liudmyla, my grandmother Olha, my partner Stelios and his wonderful family Debby, Manolis, Antonetta and Herman. Mom, thank you for being my best friend and the first and most vocal advocate of my personal and professional development. Granny, thank you for always having my back and encouraging me to be a better person. Thank you, my dear family, for being my stronghold and my biggest inspiration even in the time like this. Stelios, thank you for being there for me in good or bad, you stimulate me to always move forward and be open to new possibilities. In the past years you have been the loudest voice of reason for me. Debby, Manolis, Antonetta and Herman, you became my second family, thank you for your unconditional support and love.