Mathematical Modeling of Neutron Transport

FACULTY
OF APPLIED SCIENCES
UNIVERSITY
OF WEST BOHEMIA

Milan Hanuš

Department of Mathematics
University of West Bohemia, Pilsen

Thesis submitted in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy (Applied Mathematics)

Supervisor: Doc. Ing. Marek Brandner, Ph.D.

Pilsen, Czech Republic, 2014

Matematické modelování transportu neutronů

FAKULTA
APLIKOVANÝCH VĚD
ZÁPADOČESKÉ
UNIVERZITY
V PLZNI

Milan Hanuš

Katedra matematiky
Západočeská univerzita v Plzni

disertační práce
k získání akademického titulu *doktor*
v oboru *Aplikovaná matematika*

Školitel: Doc. Ing. Marek Brandner, Ph.D.

Plzeň, 2014

Declaration

I hereby declare that this doctoral thesis is my own work, unless clearly stated otherwise.

. . . . . .
Milan Hanuš

Abstract

The subject of this work is computational modeling of neutron transport relevant to economical and safe operation of nuclear facilities. The general mathematical model of neutron transport is provided by the linear Boltzmann's transport equation and the thesis begins with its precise mathematical formulation and presentation of known conditions for its well-posedness.

In the following part, we study approximation methods for the transport equation, starting with the classical discretization of energetic dependence and followed by the review of two most widely used methods for approximating directional dependence (the S_N and P_N methods). While these methods are usually presented independently of each other, we show that they can be put into a single framework of Hilbert space projection techniques. This fact is then used in conjunction with the results of the first part to rigorously prove rotational invariance of the P_N equations and to analyze convergence of the basic iterative scheme for solving the S_N equations. This part of the thesis is concluded by the description of a finite element method for the final discretization of spatial dependence and a discussion of the solution of the resulting system of algebraic equations.

The main new results are contained in the following two chapters focusing on the simplified P_N approximation, which is a computationally more convenient albeit not as mathematically well-founded variant of the P_N approximation. We prove well-posedness of the weak form of the SP_{3-7} equations and present a new way of deriving the equations from an alternative set to the P_N equations, obtained from special linear combination of spherical harmonics – the so-called Maxwell-Cartesian spherical harmonics, hence the abbreviation MCP_N approximation. We explicitly show how the MCP_3 equations may be transformed to the SP_3 equations.

The final part of the thesis contains numerical examples of the S_N and *hp*-adaptive SP_N calculations using a neutronics framework that has been implemented by the author to the *hp*-adaptive finite element

library Hermes2D. The SP_1 (or diffusion) model also serves as a basis of a real-world reactor calculation suite co-developed by the author for the purposes of "Project TA01020352 – Increasing utilization of nuclear fuel through optimization of an inner fuel cycle and calculation of neutron-physics characteristics of nuclear reactor cores". An example benchmark used to test the code concludes the thesis.

Keywords:

Neutron transport, Boltzmann equation, reactor criticality, multigroup approximation, PN approximation, simplified PN approximation, discrete ordinates, ray effects, source iteration, angular quadrature, projection, diffusion, spherical harmonics, well-posedness, weak formulation, Maxwell-Cartesian spherical harmonics, tensor, MCPN approximation, finite elements, discontinuous Galerkin method, hp-adaptivity, algebraic multigrid, Hermes2D, Dolfin.

Abstrakt

Práce se zabývá matematickým a numerickým modelováním transportu neutronů, se zaměřením na výpočty neutronových charakteristik jaderných reaktorů. Obecný matematický model transportu neutronů je reprezentován lineární Boltzmannovou transportní rovnicí. Práce začíná její přesnou matematickou formulaci a přehledem výsledků týkajících se její řešitelnosti ve druhé kapitole. Následující kapitoly jsou zaměřeny na přibližné metody řešení této rovnice.

Po stručném popisu klasické diskretizace energetické závislosti je hlavní část třetí kapitoly věnována aproximaci směrové závislosti pomocí dvou stěžejních metod – metody diskrétních ordinát (S_N) a metody sférických harmonických funkcí (P_N). Zatímco obvykle jsou tyto metody formulovány nezávisle, v práci je ukázáno, jak je lze obě popsat pomocí jednotného rámce jako projekci na podprostor Hilbertova prostoru funkcí definovaných na sféře. Této skutečnosti je posléze využito při důkazu rotační invariantnosti P_N rovnic a při konvergenční analýze základní iterační metody pro řešení S_N soustavy. Třetí kapitola je zakončena popisem aplikace metody konečných prvků na finální diskretizaci prostorové závislosti.

Hlavní nové výsledky této práce se týkají metody zjednodušených sférických harmonických funkcí (SP_N), jež představuje výpočetně efektivní aproximaci metody P_N. Ve čtvrté kapitole je standardním způsobem odvozena slabá formulace SP_N rovnic a dokázána její korektnost pro $N = 3, 5, 7$. V páté kapitole je pak odvozena nová soustava parciálních diferenciálních rovnic odpovídající P_N aproximaci (MCP_N aproximace). Na příkladu MCP_3 aproximace je ukázáno, jak lze využít tenzorovou strukturu těchto rovnic k transformaci na soustavu ekvivalentní s SP_3 aproximací.

V šesté kapitole je popsána implementace S_N a SP_N aproximací do knihovny Hermes2D a na několika příkladech ukázány základní vlastnosti těchto aproximací. Speciální pozornost je věnována implementaci nespojité Galerkinovy metody (pro S_N aproximaci)a modifikaci standardního indikátoru chyby pro *hp*-adaptivitu v Hermes2D pro

SP_N aproximaci. Práce je ukončena ukázkou řešení standardního 3D benchmarku pomocí mnohagrupového difúzního kódu, který autor na základě zkušeností s vývojem neutronických modulů v knihovně Hermes2D vyvinul pro účely projektu "TA01020352 – Zvýšení využití jaderného paliva pomocí optimalizace vnitřního palivového cyklu a výpočtu neutronově-fyzikálních charakt. aktivních zón jaderných reaktorů".

Klíčová slova:

Transport neutronů, Boltzmannova rovnice, kritičnost reaktoru, vícegrupová aproximace, PN aproximace, zjednodušená PN aproximace, metoda diskrétních směrů, iterace zdroje, ray efekt, směrová kvadratura, projekce, difúze, sférické harmonické funkce, korektnost, slabá formulace, Maxwellovy kartézské sférické harmonické funkce, tenzor, MCPN aproximace, konečné prvky, nespojitá Galerkinova metoda, hp-adaptivita, algebraická metoda více sítí, Hermes2D, Dolfin.

Acknowledgements

I would like to thank my supervisor Marek Brandner, for all his support and guidance throughout the course of my Ph.D. studies. Many thanks belong to all members of our TACR-PAMG team, especially to the "Himalaya Expedition group" Hanka, Roman and Zbyňák, who influenced my thinking in so many ways both direct and indirect. In particular, I would like to thank Roman for explaining to me many of his great ideas (not only about traveling or homebrewing). Having said so, I cannot fail to mention how Hanka took care of us when we were so deeply immersed in discussing these ideas, for which I am utterly grateful.

I would also like to thank all the great people whom I had an honor to work with during my Ph.D. studies and were not mentioned above. Namely to Lukáš Korous, with whom I spent so many productive days working on Hermes, Pavel Šolín for inviting me to Nevada to work on Hermes and him and his lovely wife Dáša for letting me stay in their house during that visit, and Vyacheslav Zimin for showing me and Roman the meaning of Russian hospitality during our research visit to the International Science & Technology Center in Moscow. Special thanks go to Ryan McClarren from the Texas A&M University in College Station who made it possible for me to meet the transport theory experts (including himself) and made each of my visits a smooth and enjoyable experience.

Last but certainly not least, I would like to express my deep gratitude to my family, who have always supported me in any way they could.

I acknowledge the financial support of TAČR (Technologická Agentura České Republiky) grant TA01020352 and Department of the Navy Grant N62909-11-1-7032 during the preparation of this work.

Contents

List of Figures

List of Tables

1

Introduction

Computer simulation of radiative transfer of energy is an important task in many engineering and research areas, as diverse as biomedicine, astrophysics, optics or nuclear engineering. In nuclear engineering, the area of primary interest in this work, there are two main goals of computer modeling of radiative transfer. The first is to simulate short-term transient behavior of nuclear devices under given initial conditions such as geometry and material configuration. The second is to determine under which conditions such devices (in this case typically nuclear reactor cores) will be capable of long-term, stable operation satisfying certain safety, technical and economical limitations, with only a minimal human intervention. Repeated calculations of the second type form the basis for designing new nuclear reactors or optimizing fuel reloading of existing ones. Optimization of fuel reloading schemes for nuclear reactors is the topic of a major research and development project investigated at author's department[1]. Author's participation in this project during the course of his doctoral studies involved the development of a neutron-physical calculation module that could be employed by the overall optimization suite to evaluate fitness of its candidate configurations. This fact largely influenced the choice of mathematical models and numerical methods studied in this thesis.

[1]Project TA01020352 – Increasing utilization of nuclear fuel through optimization of an inner fuel cycle and calculation of neutron-physics characteristics of nuclear reactor cores. Principal investigators: R. Čada (University of West Bohemia) and J. Rataj (Czech Technical University).

1. INTRODUCTION

The most accurate mathematical model of the physical laws governing radiative transfer mediated by mutually non-interacting particles is acknowledged to be the linear Boltzmann transport equation. In nuclear reactors, the effects of neutron-induced reactions dominate those caused by other types of particles and we will therefore consider the transport equation for neutrons in this thesis, even though it has the same form for other types of non-charged particles, such as photons. While the short-term transient simulations require accurate solution methods for the time-dependent transport equation, a quasi-steady state solution (a sequence of steady state calculations) is generally sufficient to capture slow changes in core configuration and material characteristics during its long-term stable operation. Author's work focus on the latter application domain further narrows scope of this thesis to the *steady state neutron transport equation*, shortly NTE. It is worth recalling, however, that many common numerical methods for solving transport problems involve repeated execution of methods designed for steady-state problems.

Mathematical modelling of neutron transport

We will introduce the steady state NTE in Chap. 2 as an integro-differential equation with 6 independent variables (three characterizing position of neutrons, two their streaming direction and one their energy) and review its theoretical properties. The high dimensionality of the equation requires either a direct particle simulation and use of statistical methods for obtaining the required physical quantities (the Monte Carlo approach) or a deterministic approach involving multiple discretizations. As the second approach is still preferable in terms of overall efficiency, we choose it as a basis for our research and study classical discretization methods for the NTE in Chap. 3.

We will focus on two widely used methods of this category – the *method of spherical harmonics*, abbreviated P_N and the *method of discrete ordinates*, abbreviated S_N. Both these methods can be viewed as projections of the NTE onto a particular Hilbert subspace of $L^2(\mathcal{S}_2)$ – the space of square integrable functions of the directional variables. This is the way how the P_N method is usually presented, but it is not immediately obvious in the S_N case (this will be addressed in Sec. 3.4.2). This fact will be used to study numerical behavior of the

methods by translating properties of the continuous NTE. As a first application, we will give a proof of rotational invariance property of the P_N approximation. This is a well known fact preventing the undesirable "ray effects" (Sec. 3.4.1.1) of the rotationally non-invariant S_N approximation, of which we however couldn't find a formal proof in available literature. As a second application, we will analyze convergence of a classical iterative method for the S_N approximation (Sec. 3.4.3) by direct application of a Banach fixed-point argument proved for the continuous NTE in [43].

The MCP_N approximation and its relation to the SP_N approximation In Chap. 5, we will derive a new set of equations equivalent to the original P_N set. The derivation starts by choosing an alternative approximation basis, composed of special linear combinations of the original basis used in the P_N approximation. These new basis functions (the Maxwell-Cartesian surface spherical harmonics introduced in [7]) have a clear tensorial structure formally resembling that of Legendre polynomials and lead to a set of equations resembling the 1D P_N equations. We call this set the MCP_N approximation ("Maxwell-Cartesian P_N" approximation) and use its structure to uncover its connection to another traditional approximation of neutron transport – the *simplified spherical harmonic method*, or SP_N, in the second part of Chap. 5.

The SP_N method (particularly the SP_3) already simplifies the NTE to the extent that it is applicable to day-by-day whole-core calculations on usual workstations with a few computational cores or small-scale parallel machines with tens to a few hundred cores, which are the typical machines available to nuclear engineering companies[2]. As is well known and will be recalled in Chap. 4, when the SP_3 method is applied to the typical reactor core calculations, its solution captures most of the features of the true solution of the NTE. Combined with its efficiency that allows this method to be used "off the desk" (without the need of submitting the job to some supercomputing center, waiting for it to come

[2]from personal experience of the author coming out of the long-term collaboration with the Czech nuclear engineering company Škoda JS led by author's colleague R. Kužel; more generally, see the discussion in [102, Sec. 2.4]

to the front of an execution queue and gathering the results) makes it attractive for physicists to quickly test their empirical approximations used throughout their production code, which is usually based on the most restricting transport approximation – the diffusion approximation.

Finite element framework for 2D neutron diffusion/transport

The neutron diffusion approximation, whereby the NTE is reduced to a second-order elliptic PDE (or, when energy dependence is taken into account implicitly, a weakly coupled non-symmetric system of second-order PDEs with positive-definite symmetric part – the so-called *multigroup neutron diffusion approximation*), also forms the basis of the neutronics module to be used in the above mentioned core loading optimization code. However, to obtain the final discrete algebraic system of equations, it uses the finite element method. This distinguishes it from the majority of other codes used for similar purposes, which are usually based on the so-called *nodal method*[3].

Generally speaking, a nodal method is a coarse mesh finite volume method iteratively combined with fine-level correction steps, specially tailored to the neutron diffusion (or recently SP_N) model and reactor core domain (i.e., typically, coarse level cells correspond to real fuel assemblies and the correction consists of analytic solution of the diffusion equation in a geometrically simple homogeneous region). Its advantage is the speed and overall efficiency, but it is greatly limited in geometrical flexibility. Because of the way the standard nodal equations are derived, it also requires the *homogenization* procedure to represent each coarse cell by a single set of material coefficients (or in more modern nodal methods by coefficients with a pre-specified polynomial variation) and the corresponding *de-homogenization* procedure to reconstruct the fine structure of the solution needed for further computations (where the latter, in particular, is difficult to formulate in general cases). Moreover, the convergence and stability of the method is, to the knowledge of the author, not very well understood ([123]).

This motivated the study of the feasibility of solving whole-core neutronics problems by the finite element method, which does not suffer from these issues.

[3]See e.g. [34, 60, 76] for the specific application area of core reloading optimization; some other nodal codes widely employed in various whole-core calculations are tabulated in [69].

The basic principle underlying the nodal method, i.e. comparing two solutions with different accuracy to provide a more accurate one, has led the author to the open-source finite element C++ library Hermes2D [44] , which uses the same principle to drive its advanced *hp-adaptivity* procedure [89]. Combined with its unique way of assembling coupled systems of PDEs [92], the Hermes2D library has proved to be well-suited for serving as a basis for testing the neutronic approximations described in the first part of the thesis.

The author has also participated in the development of the core library; the main contributions to the Hermes project involved:

- development of an interface for various existing sparse, direct and iterative algebraic solvers (which also required reworking the CMake build system of Hermes),
- development of a multigroup neutron diffusion framework, simplifying and unifying the formulation of multiregion, multigroup neutron diffusion problems within Hermes2D,
- development of the discontinuous Galerkin framework (together with L. Korous, the main developer of Hermes at present time) and
- extension of the h-adaptivity capabilities by the standard a-posteriori error estimation for elliptic problems (which involves solution jumps over element interfaces and thus uses elements of the discontinuous Galerkin framework).

More details about the neutronics modules for Hermes2D will be given in Chap. 6, based on the abstract weak formulation of the multigroup diffusion approximation from Chap. 4. An extension for the SP_N approximation will also be discussed, including a modification of the standard error indicator used in Hermes2D to guide the *hp*-adaptivity process. While well-posedness of the weak form of the multigroup diffusion approximation has been proved in [29, Chap. VII] or [13], we could not find a formal proof for the SP_N case and hence provide one in Sec. 4.3.

To assess the benefits of using the SP_N model over the simpler diffusion model, the author also implemented (still on top of the neutronics framework) a discontinuous Galerkin discretization of the *discrete ordinates* approximation of the

transport equation (the S_N method, studied in Section 3.4). Unlike the SP_N approximation, this approximation, theoretically as $N \to \infty$, converges to the true solution of the NTE in general multidimensional, multiregion domains. Combined with the automatic, problem independent spatial adaptivity capabilities provided by Hermes and its multimesh assembling strategy, the author expects that this implementation can serve in future as a first step for exploring adaptive solutions to more difficult transport problems not covered by the SP_N model.

3D coupled neutron-physical finite element code based on the multigroup diffusion approximation

For the purposes of the research project "Project TA01020352 – Increasing utilization of nuclear fuel through optimization of an inner fuel cycle and calculation of neutron-physics characteristics of nuclear reactor cores" (cf. the footnote on pg. 1), a 3D neutron diffusion solver was needed. Using the experience with the Hermes2D library, the author also developed a multigroup neutron diffusion solver within the FEniCS/Dolfin framework ([77, 78]). It features distributed (MPI) assembly of the multigroup neutron diffusion problem and solution of the obtained algebraic problem using the well-established PETSc/SLEPc solvers ([10, 57]) wrapped by FEniCS. This can be repeated in a feedback loop, in which the computed neutron flux directly influences thermal/hydraulic properties of the core, the change of which in turn leads to a change of coefficients in the diffusion equations. On top of that loop, another loop representing fuel burnup can be executed. At this point, the author would like to acknowledge the work of his colleagues – R. Kužel (the coordinator of the whole effort and also the author of a GPU eigensolver module), J. Egermaier and H. Kopincová (who implemented the thermal/hydraulics module) and Z. Vastl (who generated the meshes for the benchmarks on which the module has been tested).

Thorough description of the coupled code system is beyond the scope of this thesis. To give at least the glimpse of the scale of the problems that are solvable by the code, results of a selected benchmark conclude Chap. 6.

2

Mathematical model of neutron transport

The steady state neutron transport equation is a mathematical representation of balance between neutron gains and losses within a given macroscopic domain $\mathcal{D} \subset \mathbb{R}^3$. Let us consider the equation in its integro-differential form with given neutron source function q:

$$\begin{aligned}\Big[\boldsymbol{\Omega} \cdot \nabla + \sigma_t(\mathbf{r}, \boldsymbol{\Omega}, E)\Big]\psi(\mathbf{r}, \boldsymbol{\Omega}, E) = \\ = \int_{E_{\min}}^{E_{\max}} \int_{\mathcal{S}_2} \kappa(\mathbf{r}, \boldsymbol{\Omega} \leftarrow \boldsymbol{\Omega}', E \leftarrow E')\psi(\mathbf{r}, \boldsymbol{\Omega}', E', t)\, \mathrm{d}\boldsymbol{\Omega}'\, \mathrm{d}E' + q(\mathbf{r}, \boldsymbol{\Omega}, E).\end{aligned} \tag{2.1}$$

Function σ_t groups all reactions that result in a loss of neutron, while κ represents reactions that introduce neutrons into direction $\boldsymbol{\Omega}$ and energy E by, e.g., scattering from direction $\boldsymbol{\Omega}'$, slowing down (or accelerating) from higher (lower) energies E' or releasing new neutrons from fissioned nuclei. They are given by material composition of the domain and we will return to their more detailed description, as well as to boundary conditions for (2.1), in Section 2.2.

Solution of eq. (2.1), the *angular neutron flux density* ψ – is a function of the following independent variables, which define the neutron phase space:

- $\mathbf{r} = (x, y, z)$ represents the spatial distribution of neutrons,
- $\boldsymbol{\Omega}$ represents the angular distribution of neutrons on a unit sphere $\mathcal{S}_2$, i.e. their streaming direction ($\boldsymbol{\Omega} \in \mathbb{R}^3$, $\|\boldsymbol{\Omega}\| = 1$);

- $E \in [E_{\min}, E_{\max}]$ is the kinetic energy of neutrons.

REMARK 1. The phase space could be also defined in terms of the velocity vector $\mathbf{v}$ and speed $v = \|\mathbf{v}\| = \sqrt{2E/m}$ (m being neutron mass) instead of $\mathbf{\Omega}$ and E. This form appears to be preferred in analytical works, while our choice is more often used in practical numerical calculations. Corresponding changes in the formulation of the NTE are explicitly given e.g. in [30, Chap. XXI, eqns. (1.1) and (1.2)]. For further use, we will just note that we can write $\mathbf{\Omega} = \mathbf{v}/v$ with $\mathbf{v} \in \mathbb{R}^3$.

REMARK 2. In this macroscopic description, we should always consider beams of neutrons with the same average properties in differential elements around $\mathbf{r}$, $\mathbf{\Omega}$, E. For simplicity, we will refer to them as to single neutrons with particular position, energy or direction (and occasionally call them $(\mathbf{r}, \mathbf{\Omega}, E)$-neutrons) and we also omit the "density" nomenclature of repeatedly used quantities – e.g., we will henceforth call ψ just *angular neutron flux*.

REMARK 3. We keep in mind that eq. (2.1) is a consequence of applying Gauss divergence theorem to the fundamental integral neutron balance over an arbitrary bounded subdomain of the phase space (which assumes differentiable angular neutron flux).

2.1 Neutron phase space

Let us assume that $\mathcal{D}$ is a domain bounded by a piecewise smooth boundary $\partial\mathcal{D}$, which is oriented at almost every point $\mathbf{r} \in \partial\mathcal{D}$ (a.e. in $\partial\mathcal{D}$) by its unit outward normal field $\mathbf{n}(\mathbf{r})$. Then we may formally define the neutron phase space

$$X \coloneqq \{(\mathbf{r}, \mathbf{\Omega}, E) : \ \mathbf{r} \in \mathcal{D} \subset \mathbb{R}^3, \mathbf{\Omega} \in \mathcal{S}_2, E \in [E_{\min}, E_{\max}]\}$$

together with its outflow and inflow boundary subsets, respectively:

$$\partial X^{\pm} \coloneqq \big\{(\mathbf{r}, \mathbf{\Omega}, E) \in \partial\mathcal{D} \times \mathcal{S}_2 \times [E_{\min}, E_{\max}], \text{ s.t. } \mathbf{\Omega} \cdot \mathbf{n}(\mathbf{r}) \gtrless 0\big\}.$$

The whole boundary ∂X can be written as

$$\partial X = \partial X^+ \cup \partial X^- \cup \partial X^0$$

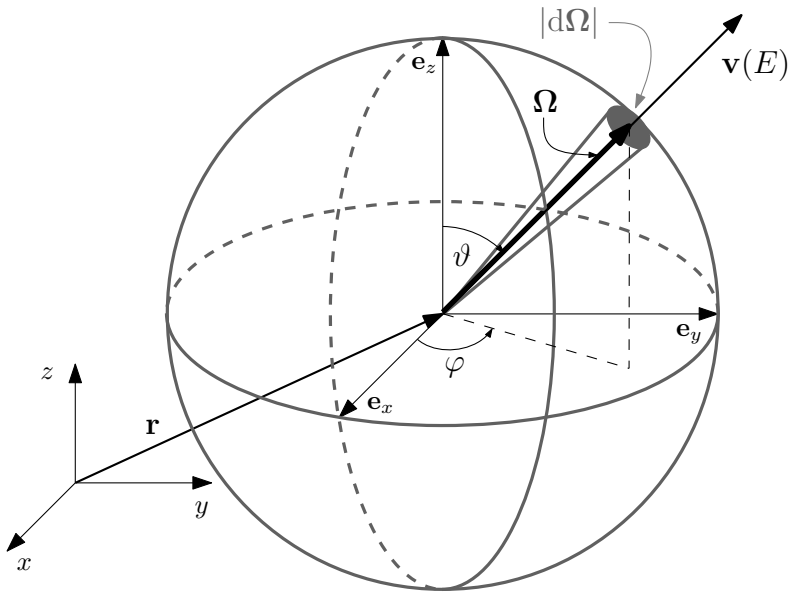


Figure 2.1: Phase space of neutrons

where ∂X^0 (boundary subset tangential to the flow), so that $\overline{X} = X \cup \partial X$. The product Lebesgue measure

$$\mathrm{d}x \;= \mathrm{d}\mu(X) \;= \mathrm{d}\mu(V \times \mathcal{S}_2 \times [E_{\min}, E_{\max}]) \;= \mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{\Omega}\,\mathrm{d}E \tag{2.2}$$

is used when integrating over X, while the boundary measure

$$\mathrm{d}\xi \;= |\mathbf{\Omega} \cdot \mathbf{n}|\,\mathrm{d}S\,\mathrm{d}\mathbf{\Omega}\,\mathrm{d}E\,, \tag{2.3}$$

is used when integrating over $\partial X^{\pm}$, where $S = \mu(\partial\mathcal{D})$. Note that because of the assumed regularity of the boundary, ∂X^0 is a closed subset of ∂X of $\mathrm{d}\xi$-measure zero ([30, Chap. XXI, Sec. 2.2]). This will allow us to decompose spaces of measurable functions defined on ∂X into a direct sum of subspaces of measurable functions defined on ∂X^+ and ∂X^-, respectively. When referring to physical units, we will consider the length scale of $\mathcal{D}$ in centimeters.

Since the direction vectors are confined to the sphere, we can express the three Cartesian components of $\mathbf{\Omega}$ by only two spherical coordinates $\vartheta \in [0, \pi]$ and $\varphi \in [0, 2\pi)$:el

$$\mathbf{\Omega} = \begin{bmatrix} \Omega_x \\ \Omega_y \\ \Omega_z \end{bmatrix} = \begin{bmatrix} \sin\vartheta\cos\varphi \\ \sin\vartheta\sin\varphi \\ \cos\vartheta \end{bmatrix}$$

(see Fig. 2.2). To transform integrals with respect to $\mathrm{d}\mathbf{\Omega}$ into double integrals

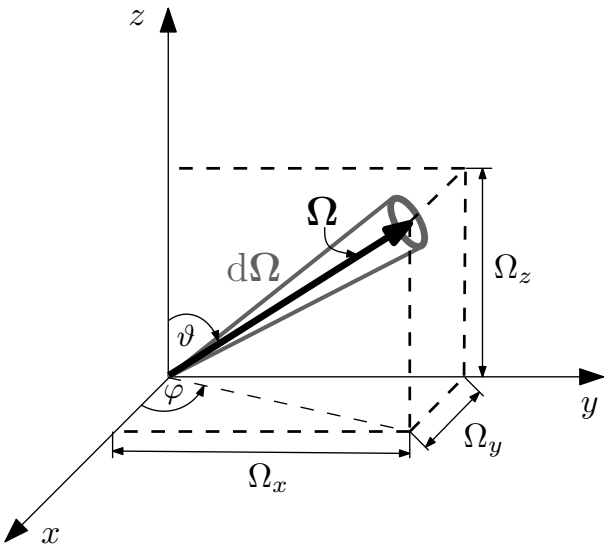


Figure 2.2: Cartesian coordinate system

with respect to ϑ and φ, note that the solid angle $\mathrm{d}\mathbf{\Omega}$ subtended at the center of $\mathcal{S}_2$ by the spherical differential element $|\mathrm{d}\mathbf{\Omega}\,|$ can be written as:

$$\mathrm{d}\mathbf{\Omega} = \frac{|\mathrm{d}\mathbf{\Omega}\,|}{r^2} = \frac{r^2 \sin\vartheta \mathrm{d}\vartheta\,\mathrm{d}\varphi}{r^2} = \sin\vartheta \mathrm{d}\vartheta\,\mathrm{d}\varphi$$

(see Fig. 2.3). We will also need to integrate functions that depend on the cosine of the angle between two directions $\mathbf{\Omega}$ and $\mathbf{\Omega}'$. We shall denote this angle and its cosine by ϑ_0 and μ_0, respectively (see Fig. A.1 in appendix for geometrical interpretation). Then

$$\mu_0 \equiv \cos\vartheta_0 = \mathbf{\Omega}\cdot\mathbf{\Omega}'$$

and

$$\begin{aligned}\int_{\mathcal{S}_2} f(\mathbf{\Omega}\cdot\mathbf{\Omega}')\,\mathrm{d}\mathbf{\Omega}' &= \int_0^{2\pi}\int_0^{\pi} f(\cos\vartheta_0)\sin\vartheta_0 \mathrm{d}\vartheta_0 = 2\pi\int_{-1}^{1} f(\mu_0)\mathrm{d}\mu_0 \\ &= \int_{\mathcal{S}_2} f(\mathbf{\Omega}'\cdot\mathbf{\Omega})\,\mathrm{d}\mathbf{\Omega}\end{aligned} \tag{2.4}$$

Notice that the result depends on neither $\mathbf{\Omega}$ nor $\mathbf{\Omega}'$.

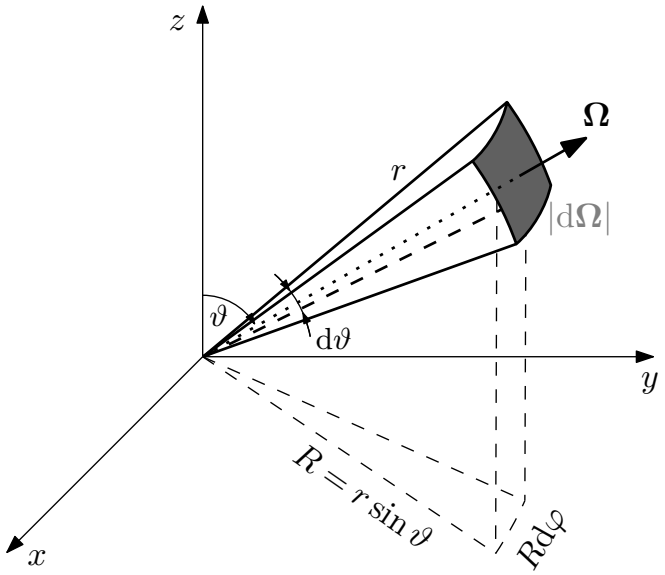


Figure 2.3: Schematic view of the solid angle of directions

2.2 Steady state neutron transport in isotropic bounded domain

In most practical cases, we can assume that the medium in which we study neutron transport is isotropic. The first consequence of this assumption is that

$$\sigma_t(\mathbf{r}, \mathbf{\Omega}, E)\psi(\mathbf{r}, \mathbf{\Omega}, E) \equiv \sigma_t(\mathbf{r}, E)\psi(\mathbf{r}, \mathbf{\Omega}, E)$$

for all $\mathbf{r}$, $\mathbf{\Omega}$, E. The second is that reactions that change the direction of neutrons from $\mathbf{\Omega}'$ to $\mathbf{\Omega}$ are invariant under rotation of the coordinate system and are thus completely determined by the cosine of the two vectors:

$$\kappa(\cdot, \mathbf{\Omega} \leftarrow \mathbf{\Omega}', \cdot) \equiv \kappa(\cdot, \mathbf{\Omega} \cdot \mathbf{\Omega}', \cdot). \tag{2.5}$$

The steady state NTE (2.1) in this regime reads

$$\begin{aligned} &\mathbf{\Omega} \cdot \nabla\psi(\mathbf{r}, \mathbf{\Omega}, E) + \sigma_t(\mathbf{r}, E)\psi(\mathbf{r}, \mathbf{\Omega}, E) = \\ &= \int_{E_{\min}}^{E_{\max}} \int_{\mathcal{S}_2} \kappa(\mathbf{r}, \mathbf{\Omega} \cdot \mathbf{\Omega}', E \leftarrow E')\psi(\mathbf{r}, \mathbf{\Omega}', E') \,\mathrm{d}\mathbf{\Omega}' \,\mathrm{d}E' + q(\mathbf{r}, \mathbf{\Omega}, E) \end{aligned} \tag{2.6}$$

in X, complemented by specified angular flux distribution at ∂X^-. The two prototypical inflow boundary conditions are:

- incoming angular neutron flux

$$\psi|_{\partial X^-} = \psi_{\text{in}} \tag{2.7}$$

($\psi_{\text{in}} \equiv 0$ corresponds to vacuum in $\mathbb{R}^3 \setminus \overline{V}$, which is a common assumption in nuclear reactor modeling),

- albedo boundary reflection

$$\psi(\mathbf{r}, \mathbf{\Omega}, E) = \beta(\mathbf{r})\psi(\mathbf{r}, \mathbf{\Omega}_R, E), \quad (\mathbf{r}, \mathbf{\Omega}, E) \in \partial X^-, \quad \mathbf{\Omega}_R = \mathbf{\Omega} - 2\mathbf{n}(\mathbf{\Omega} \cdot \mathbf{n}) \tag{2.8}$$

where $\mathbf{\Omega}$ is the reflection of $\mathbf{\Omega}_R$ about the boundary plane. For $\beta \equiv 1$, this corresponds to complete specular reflection and is used to model planes of symmetry, while for $\beta \equiv 0$, we recover the vacuum condition from above. Intermediate values mean that a fraction of neutrons leaving the domain in direction $\mathbf{\Omega}_R$ are returned back in direction $\mathbf{\Omega}$, which is commonly used to model reactor reflectors. We thus assume $0 \leq \beta \leq 1$.

Remark 4. The albedo coefficient β may in general vary with the reflector properties and should also capture redistribution of the reflected neutrons within the phase space due to their diffusion through the reflector. A general treatment of albedo condition is given in [100] (see also [101]), where an integral albedo operator B_β is introduced, such that

$$\psi(x)\big|_{\partial X^-} = (B_\beta \psi)(x) = \int_{\partial X^+} \mathrm{d}\xi\, '\beta(x \leftarrow x')\psi(x'), \tag{2.9}$$

where $x = (\mathbf{r}, \mathbf{\Omega}, E)$ and $x' = (\mathbf{r}', \mathbf{\Omega}', E')$.

For formal description of other types of boundary conditions, we refer to [100] or [4, Sec. 1.3].

A physically plausible solution of the NTE should be moreover non-negative throughout $\mathcal{D}$ and continuous along any direction $\mathbf{\Omega}$, i.e. $\psi(\mathbf{r} + s\mathbf{\Omega}, \mathbf{\Omega}, E)$ is a continuous function of s for any $\mathbf{r}$, $\mathbf{\Omega}$, E. Note that $\psi(\mathbf{r} + s\mathbf{\Omega}', \mathbf{\Omega}, E)$ *may* be discontinuous as a function of position when $\mathbf{\Omega}' \neq \mathbf{\Omega}$.

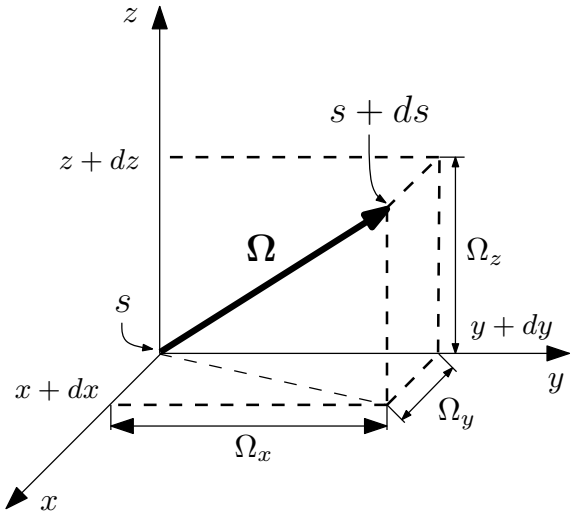


Figure 2.4: Characteristic direction in the Cartesian coordinate system

2.2.1 Advection term

In Cartesian coordinate system (that we will exclusively consider in this thesis),

$$\boldsymbol{\Omega} \cdot \nabla \psi = \boldsymbol{\Omega}_x \frac{\partial \psi}{\partial x} + \boldsymbol{\Omega}_y \frac{\partial \psi}{\partial y} + \boldsymbol{\Omega}_z \frac{\partial \psi}{\partial z} = \frac{dx}{ds}\frac{\partial \psi}{\partial x} + \frac{dy}{ds}\frac{\partial \psi}{\partial y} + \frac{dz}{ds}\frac{\partial \psi}{\partial z} = \frac{d\psi}{ds},$$

where $s \in I \subset \mathbb{R}$ parametrizes the path traveled by the neutron along the direction $\boldsymbol{\Omega}$ (the *characteristic*, see Fig. 2.4). Assuming now for simplicity that the integral term on the right of (2.6) is absorbed in the source term q, we may invert the differential operator on the left of (2.6) by integration along these characteristics and obtain an integral formulation of the neutron transport equation:

$$\psi(\mathbf{r}, \boldsymbol{\Omega}) = \psi(\mathbf{r}_0, \boldsymbol{\Omega}) e^{-\tau(\mathbf{r}, \mathbf{r}_0)} + \int_0^{s_0} q(\mathbf{r}', \boldsymbol{\Omega}) e^{-\tau(\mathbf{r}, \mathbf{r}')} \, \mathrm{d}s' \tag{2.10}$$

where

$$\mathbf{r}' = \mathbf{r} - s'\boldsymbol{\Omega}, \quad \mathbf{r}_0 = \mathbf{r} - s_0\boldsymbol{\Omega}$$

$$\tau(\mathbf{r}, \mathbf{r}') = \tau(\mathbf{r}, \mathbf{r} - s'\boldsymbol{\Omega}) = \int_0^{s'} \sigma_t(\mathbf{r} - s''\boldsymbol{\Omega}) \, \mathrm{d}s'' \tag{2.11}$$

In reference to Fig. 2.5 we can interpret the first term on the right of (2.10) as the number of neutrons moving in the direction $\boldsymbol{\Omega}$ that entered the given volume at

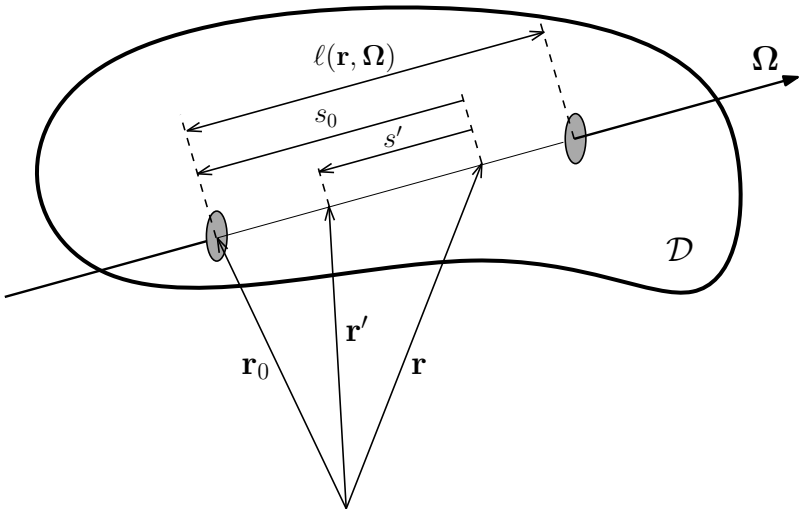


Figure 2.5: Illustration for the solution on a characteristic

$\mathbf{r}_0$ and reached point $\mathbf{r}$ without collision, whereas the second term as the number of neutrons introduced into the characteristic direction by sources between $\mathbf{r}_0$ and $\mathbf{r}$ and reaching $\mathbf{r}$ without collision. The *optical path length* τ represents the probability of collision between r' and r. This integral form is of both theoretical and practical value, as we will see later in sections 2.2.5 and 3.2.1.

2.2.2 Collision terms

The kernel of the integral operator on the right-hand side of (2.6) characterizes the mean number of ($\mathbf{\Omega}$, E)-neutrons coming out of a collision of ($\mathbf{\Omega}'$, E')-neutrons with nuclei at $\mathbf{r}$ (or more precisely in a differential element around $\mathbf{r}$). Such collisions can either just change the direction and energy of the inducing neutrons (elastic scattering) or cause absorption of the neutrons followed by release of new ones in the considered direction and energy range (fission), or both (inelastic scattering). This categorization motivates the splitting

$$\kappa(\mathbf{r}, \mathbf{\Omega} \cdot \mathbf{\Omega}', E \leftarrow E') = \eta\sigma_s(\mathbf{r}, \mathbf{\Omega} \cdot \mathbf{\Omega}', E \leftarrow E') + \nu\sigma_f(\mathbf{r}, \mathbf{\Omega} \cdot \mathbf{\Omega}', E \leftarrow E'). \quad (2.12)$$

where the two components are the so-called *double-differential macroscopic cross-section* for scattering and fission, respectively. The *scattering* and *fission yield*

η and ν, respectively, have the meaning of expected number of neutrons coming out of the scattering event (= 1 in case of elastic scattering, might be > 1 when inelastic scattering takes place) and the fission event, respectively, per one inducing neutron.

Ordinary macroscopic cross-sections $\sigma_s(\mathbf{r}, E)$, $\sigma_f(\mathbf{r}, E)$ are then introduced to characterize the total probability that $(\mathbf{\Omega}, E)$ neutrons undergo collisions of the above type irrespective of the outgoing (primed) direction and energy[1], i.e.

$$\begin{aligned} \eta\sigma_s(\mathbf{r}, E) &= \int_{E_{\min}}^{E_{\max}} \int_{\mathcal{S}_2} \eta\sigma_s(\mathbf{r}, \mathbf{\Omega}' \cdot \mathbf{\Omega}, E' \leftarrow E)\, \mathrm{d}\mathbf{\Omega}'\, \mathrm{d}E'\,, \\ \nu\sigma_f(\mathbf{r}, E) &= \int_{E_{\min}}^{E_{\max}} \int_{\mathcal{S}_2} \nu\sigma_f(\mathbf{r}, \mathbf{\Omega}' \cdot \mathbf{\Omega}, E' \leftarrow E)\, \mathrm{d}\mathbf{\Omega}'\, \mathrm{d}E'\,. \end{aligned} \tag{2.13}$$

The *total macroscopic cross-section*, σ_t, characterizing the probability that neutron with energy E undergoes a collision of any type with nuclei at $\mathbf{r}$, can now be decomposed as

$$\sigma_t(\mathbf{r}, E) = \sigma_c(\mathbf{r}, E) + \sigma_f(\mathbf{r}, E) + \sigma_s(\mathbf{r}, E) \equiv \sigma_a(\mathbf{r}, E) + \sigma_s(\mathbf{r}, E) \tag{2.14}$$

where

- $\sigma_c(\mathbf{r}, E)$ is the non-productive capture cross section (resulting in no new neutrons being introduced into the system) and
- $\sigma_a(\mathbf{r}, E) = \sigma_c(\mathbf{r}, E) + \sigma_f(\mathbf{r}, E)$ is the absorption cross section.

For later use, note that fission is an isotropic process (which removes the angular dependence of the double-differential fission cross-section altogether) and the new energy distribution does not depend on energy of the inducing neutron (as it corresponds to neutrons originally bound inside the nucleus; Fig. 2.6 shows a typical shape of that function); using the second eq. (2.13), we can then write:

$$\nu\sigma_f(\mathbf{r}, \mathbf{\Omega} \cdot \mathbf{\Omega}', E \leftarrow E') \equiv \frac{\chi(E)\nu\sigma_f(\mathbf{r}, E')}{4\pi}, \quad \int_{E_{\min}}^{E_{\max}} \chi(E)\, \mathrm{d}E = 1. \tag{2.15}$$

[1]Recall that the direction $\mathbf{\Omega}$ of the incoming neutrons is irrelevant for the result as a consequence of the assumption of isotropic medium (and eq. (2.4)).

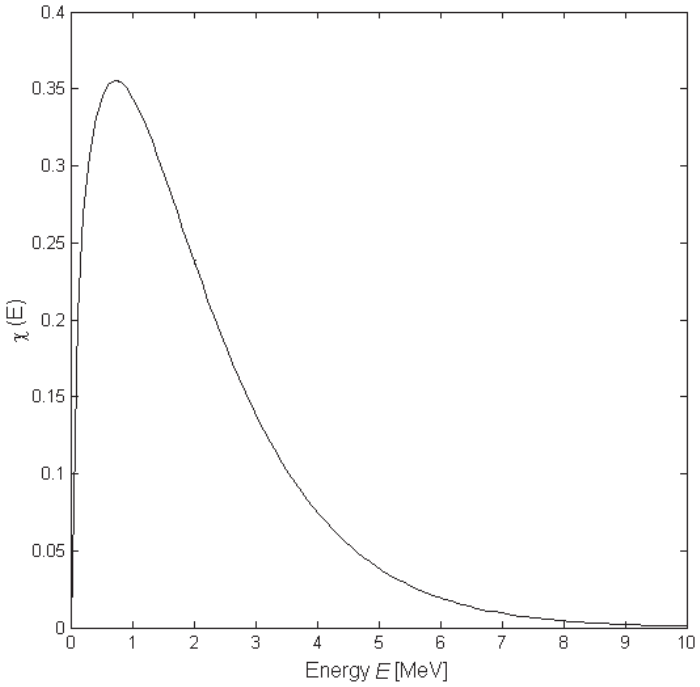


Figure 2.6: Energy spectrum of (prompt) neutrons released from fission of U235

A physically realistic assumption is that all the macroscopic cross-sections are bounded measurable functions[2], piecewise continuous in $\mathcal{D}$. The unit of macroscopic cross-sections is cm^{-1}.

2.2.3 Quantities of interest

From the solution of eq. (2.6), one can derive the following important integral quantities

- *scalar neutron flux (density)* $[\mathrm{cm}^{-2}\cdot\mathrm{s}^{-1}]$

$$\phi(\mathbf{r}, E) = \int_{\mathcal{S}_2} \psi(\mathbf{r}, \mathbf{\Omega}, E)\, \mathrm{d}\mathbf{\Omega}\,, \tag{2.16}$$

[2]with respect to product measure of type (2.2) appropriate for their particular set of arguments, or with respect to $\mathrm{d}\mu(V \times \mathcal{S}_2^2 \times [E_{\min}, E_{\max}]^2)$ in case of double-differential cross-section

- *net neutron current (density)* $[\mathrm{cm}^{-2}\cdot\mathrm{s}^{-1}]$

$$\mathrm{J}(\mathbf{r}, E) = \int_{\mathcal{S}_2} \mathbf{\Omega}\psi(\mathbf{r}, \mathbf{\Omega}, E)\,\mathrm{d}\mathbf{\Omega}\,, \tag{2.17}$$

so that integrating $\mathrm{J}(\mathbf{r}, E)\cdot\mathbf{n}(\mathbf{r})$ over a given surface gives the total number of neutrons with energy E crossing (per unit time) that surface in the direction of $\mathbf{n}$ (and allows assessing neutron conservation within given volume, recall Remark 3).

The integral

$$\int_{E_1}^{E_2} \sigma_x(\mathbf{r}, E)\phi(\mathbf{r}, E)\,\mathrm{d}E \tag{2.18}$$

represents the *reaction rate density* (per unit time) of given type ($x = t, a, f, s, c$, see (2.14)), induced by neutrons of energies in range $[E_1, E_2]$. As well as the scalar flux itself, reaction rates may be experimentally measured by various detector mechanisms, which is the reason why these quantities are more important in practical calculations than the actual solution of the NTE (angular neutron flux ψ). Of particular importance for reactor calculations is the *power density*

$$P(\mathbf{r}) = \int_{E_{\min}}^{E_{\max}} e\sigma_f(\mathbf{r}, E)\phi(\mathbf{r}, E)\,\mathrm{d}E \quad [\mathrm{W}\cdot\mathrm{cm}^{-3}] \tag{2.19}$$

where e is the energy conversion factor converting fission rate to watts.

2.2.4 Solvability of neutron transport problems

In this section, we will formulate two basic problems of neutron transport in an operator form. We will be interested in *generalized solutions* of these problems (which we call just solutions), which satisfy the equation and boundary conditions almost everywhere (a.e.) in X (or ∂X), and understand by $\mathbf{\Omega}\cdot\nabla$ the generalized directional derivative in the usual Sobolev sense. This is motivated by the low regularity that can be expected from the exact solution of the NTE – for example, even for piecewise smooth material data (cross-sections σ_x) and sources q, the solution of the NTE is known to possibly exhibit singularities in first partial derivatives (or be discontinuous as a function of $\mathbf{\Omega}$) at surfaces of material discontinuities coinciding with characteristic curves of the left-hand side of (2.6) (that is, straight lines; see [4, Chap. 1], [117, Sec. III]).

2.2.5 Neutron transport problem with fixed sources

Let us introduce the standard spaces of Lebesgue-integrable functions (w.r.t. the product measure (2.2), resp. (2.3))

$$\begin{aligned}
L^p(X) &= \left\{ \psi \mid \|\psi\|_{L^p(X)} \coloneqq \left(\int_X |\psi(x)|^p \, \mathrm{d}x \right)^{1/p} < \infty \right\}, \quad 1 \le p < \infty, \\
L^p(\partial X^\pm) &= \left\{ \psi \mid \|\psi\|_{L^p(\partial X^\pm)} \coloneqq \left(\int_{\partial X^\pm} |\psi(x)|^p \, \mathrm{d}\xi \right)^{1/p} < \infty \right\}, \quad 1 \le p < \infty, \\
L^\infty(X) &= \{ \psi \mid \|\psi\|_{L^\infty(X)} \coloneqq \operatorname{ess\,sup}_X |\psi(x)| < \infty \} \\
L^\infty(\partial X^\pm) &= \{ \psi \mid \|\psi\|_{L^\infty(\partial X^\pm)} \coloneqq \operatorname{ess\,sup}_{\partial X^\pm} |(\boldsymbol{\Omega} \cdot \mathbf{n})\psi(x)| < \infty \}
\end{aligned} \tag{2.20}$$

Note that for the total volumetric scalar flux to be finite (as is physically expected), the solution should belong to $L^1(X)$.

We formulate the fixed source problem using the following operators:

$$\begin{aligned}
A\psi(\mathbf{r}, \boldsymbol{\Omega}, E) &= \boldsymbol{\Omega} \cdot \nabla \psi(\mathbf{r}, \boldsymbol{\Omega}, E), \\
\Sigma_t \psi(\mathbf{r}, \boldsymbol{\Omega}, E) &= \sigma_t(\mathbf{r}, E) \psi(\mathbf{r}, \boldsymbol{\Omega}, E), \\
K\psi(\mathbf{r}, \boldsymbol{\Omega}, E) &= \int_{E_{\min}}^{E_{\max}} \int_{\mathcal{S}_2} \kappa(\mathbf{r}, \boldsymbol{\Omega} \cdot \boldsymbol{\Omega}', E \leftarrow E') \psi(\mathbf{r}, \boldsymbol{\Omega}', E') \, \mathrm{d}\boldsymbol{\Omega}' \, \mathrm{d}E' \,.
\end{aligned}$$

We shall call A, Σ_t, K and $T = A + \Sigma_t - K$ the *advection*, *reaction*, *collision* and *transport* operator, respectively. All these operators are continuous; the reaction operator $\Sigma_t : L^p(X) \to L^p(X)$ is a simple multiplication operator in $L^p(X)$, self-adjoint and bounded, while the operator $K : L^p(X) \to L^p(X)$ is bounded under additional (physically justifiable) conditions (conditions (c) and/or (d) of the following theorem, depending on p), but is self-adjoint if and only if the kernel κ of K is symmetric in E and E'. With the exception of the mono-energetic case, this is generally not true (a fact to which we return again in Sec. 3.1). To further simplify notation, let us also define

$$L \coloneqq A + \Sigma_t.$$

For piecewise smooth $\partial\mathcal{D}$ and $1 \le p < \infty$, the traces $\Gamma_\pm \psi \equiv \psi|_{\partial X^\pm} \in L^p(\partial X^\pm)$ are well defined for functions $\psi \in \widehat{H}^p(X)$, where

$$\widehat{H}^p(X) = \{ \psi \mid \psi \in L^p(X), \boldsymbol{\Omega} \cdot \nabla \psi \in L^p(X) \}, \quad 1 \le p \le \infty$$

and a continuous lifting operator

$$\mathcal{G} : \psi_{\pm} \in L^p(\partial X^{\pm}) \mapsto \widehat{\psi} \in \widehat{H}^p(X)$$

such that $\widehat{\psi}|_{\partial X^{\pm}} = \psi_{\pm}$ exists ([30, Thm. 1, Appendix of §2, Chap. XXI], [11] for the case $p = \infty$). The Sobolev space of functions with bounded boundary traces is then defined as

$$H^p(X) \coloneqq \left\{\psi \in L^p(X), \boldsymbol{\Omega} \cdot \nabla\psi \in L^p(X), \|\Gamma\psi\|_{L^p(\partial X)} < \infty\right\}. \tag{2.21}$$

and $A : H^p(X) \to L^p(X)$, thus the complete transport operator $T : H^p(X) \to L^p(X)$. We note that $H^2(X)$ is a Hilbert space when equipped with the inner product

$$(\psi, \varphi)_{H^2(X)} \coloneqq (\boldsymbol{\Omega} \cdot \nabla\psi, \boldsymbol{\Omega} \cdot \nabla\varphi)_{L^2(X)} + (\psi, \varphi)_{L^2(X)} + (\psi, \varphi)_{L^2(\partial X)}$$

where

$$(\psi, \varphi)_{L^2(X)} = \int_X \psi\varphi \,\mathrm{d}x\,, \quad (\psi, \varphi)_{L^2(\partial X)} = \int_{\partial X} \psi\varphi \,\mathrm{d}\xi\,. \tag{2.22}$$

The lifting operator allows us to pick a function $\widehat{\psi} = \mathcal{G}\psi_{\text{in}} \in H^p(X)$ and convert a problem $T\psi = q$ with non-homogeneous boundary conditions (2.7) to a problem

$$T(\psi - \widehat{\psi}) = q - T\widehat{\psi} \equiv \widehat{q}$$

where trace of the new unknown function $u = \psi - \widehat{\psi}$ on ∂X^- vanishes. Final solution is then recovered as $\psi = u + \widehat{q}$. Therefore, we can focus on the case with homogeneous conditions and put $\mathrm{Dom}\,(T) = H_0^p(X)$ where

$$H_0^p(X) \coloneqq \{\psi \in H^p(X),\ \psi|_{\partial X^-} = 0\}.$$

Similar treatment of reflective or more general boundary conditions requires special trace theorems, see [30, Chap. XXI, Appendix of §2] or [4, Chap. 2].

The fixed source, steady state neutron transport problem with vacuum boundary conditions that we are going to study in this section is posed as follows

Problem 1. *For given $q \in L^p(X)$ find $\psi \in H_0^p(X) \subset L^p(X)$ such that $T\psi = q$.*

We will first consider the physically most natural $L^1(X)$ case.

2.2.5.1 $L^1(X)$ setting

Theorem 1. *Assume that*

(a) $\sigma_t \in L^\infty(X)$, $\sigma_t \geq \underline{\sigma_t} > 0$ *a.e. in* $\mathcal{D} \times [E_{min}, E_{max}]$,

(b) $\kappa \geq 0$ *a.e. in* $\mathcal{D} \times \mathcal{S}_2^2 \times [E_{min}, E_{max}]^2$,

(c) $c \leq \bar{c} < 1$ *a.e. in* $\mathcal{D} \times [E_{min}, E_{max}]$ *where*

$$c(\mathbf{r}, E) \coloneqq \frac{1}{\sigma_t(\mathbf{r}, E)} \int_{E_{min}}^{E_{max}} \int_{\mathcal{S}_2} \kappa(\mathbf{r}, \boldsymbol{\Omega}' \cdot \boldsymbol{\Omega}, E' \leftarrow E)\, dE'\, d\boldsymbol{\Omega}' \,. \tag{2.23}$$

Then Problem 1 with $p = 1$ *has a unique solution* $\psi \in H_0^1(X)$.

Proof. [30, Chap. XXI, §2, Proposition 5] □

The value c in Thm. 1 has the physical meaning of the mean (net) number of neutrons emitted (in all possible directions and energies) per a neutron with energy E (coming from any direction) colliding with a nucleus at point $\mathbf{r} \in V$. Condition (c) thus expresses the requirement that the system be *subcritical* in order for a steady solution in presence of external sources to be achieved (the notion of criticality will be formally introduced in the following subsection). Notice that (using (2.13))

$$c(\mathbf{r}, E) = \frac{\eta\sigma_s(\mathbf{r}, E) + \nu\sigma_f(\mathbf{r}, E)}{\sigma_t(\mathbf{r}, E)} \tag{2.24}$$

and is usually called *collision ratio* (or *scattering ratio* in non-fissioning domains).

In [101], Sanchez uses the inversion of the transport operator along characteristics (see 2.2.1 above) to prove existence and uniqueness of solution to the fixed source neutron transport problem in the cross-section weighted space $H_\sigma^1(X)$ for right hand sides in $L_\sigma^1(X)$ (and albedo boundary conditions of general type (2.9)).This appears to be an alternative physically natural functional setting due to the definition of reaction rate, eq. (2.18); moreover, assumption (a) may be relaxed by allowing $\sigma_t = 0$ in arbitrarily large regions (the *void regions*). Note that if we assume (a) of Thm. 1, the norms of L^1 and L_σ^1 are equivalent and the measure $\mathrm{d}\tau = \sigma_t(x)\mathrm{d}x$ associated with the space L_σ^1 represents a differential optical path length (see eq. (2.11)).

2.2.5.2 $L^\infty(X)$ setting

In [30, Chap. XXI, §2, Proposition 6], Dautray and Lions also show the existence of unique solution in $L^\infty(X)$ for $q \in L^\infty(X)$; the proof in this case is again based on the inversion of the transport operator along characteristics and requires instead of (c) the condition

(d) $d \leq \overline{d} < 1$ a.e. in $\mathcal{D} \times [E_{\min}, E_{\max}]$

where

$$d(\mathbf{r}, E) \coloneqq \frac{1}{\sigma_t(\mathbf{r}, E)} \int_{E_{\min}}^{E_{\max}} \int_{\mathcal{S}_2} \kappa(\mathbf{r}, \boldsymbol{\Omega} \cdot \boldsymbol{\Omega}', E \leftarrow E') \, \mathrm{d}E' \, \mathrm{d}\boldsymbol{\Omega}' \tag{2.25}$$

can be interpreted as the average number of neutrons emitted with energy E from collisions induced by all possible neutrons impinging on the nucleus at $\mathbf{r}$ (again, this is a reasonable condition in the subcritical state). Notice that because angular dependence of κ is only through the cosine of the collision angle (i.e., $\boldsymbol{\Omega} \cdot \boldsymbol{\Omega}'$), the outgoing direction is immaterial and assumptions (c) and (d) really represent different assumptions about just the energy transfer in collisions.

2.2.5.3 $L^2(X)$ setting and second-order forms of NTE

In general $L^p(X)$ spaces with $1 < p < \infty$, Dautray and Lions outline the proof based on the same ideas as those used in the $L^1(X)$ case (theory of monotone operators), utilizing assumptions (a-d) of Thm. 1. The case $p = 2$ is particularly important as the Hilbert space structure of $L^2(X)$ allows to use richer set of mathematical tools to formulate practical solution methods. In particular, it allows to formulate variational principles for the NTE, which are of both theoretical and practical importance.

Let us set $V = H_0^2(X)$ and let V' denote the dual space, i.e. the space of bounded linear functionals on V. For $T : V \to V'$, we may view the problem of finding $\psi \in V$ such that $T\psi = q$ as a problem posed in V' and write it in the variational (or *weak*) form: Find $\psi \in V$ such that

$$\langle T\psi, \varphi \rangle = \langle q, \varphi \rangle \quad \forall \varphi \in V, \tag{2.26}$$

where $\langle \cdot, \cdot \rangle$ denotes the duality pairing between V' and V. As usual, we will refer to the arbitrarily varying functions $\varphi \in V$ in (2.26) as to *test functions*.

This formulation allows considering more general right hand sides $q \in V' \supset L^2(X)$, but we will restrict our attention to the case represented by Problem 1 with $p = 2$, i.e. $T : V \to L^2(X)$ with $q \in L^2(X)$. Since for $L^2(X)$ the duality pairing coincides with the ordinary $L^2(X)$ inner product (2.22) and the Riesz representation theorem lets us identify $L^2(X)$ with its dual, we also identify the linear functionals $Tu, q \in [L^2(X)]'$ with their Riesz representants $Tu, q \in L^2(X)$. If we now define the bilinear and linear form

$$a(u, v) = (Tu, v)_{L^2(X)}, \quad f(v) = (q, v)_{L^2(X)},$$

we can write the weak formulation (2.26) as

Problem 2. *For given $q \in L^2(X)$ find $\psi \in V$ such that*

$$a(\psi, \varphi) = f(\varphi) \quad \forall \varphi \in L^2(X).$$

Moreover, for a sufficiently smooth boundary, the incoming boundary conditions can be imposed weakly by incorporating them into the bilinear form $a(u, v)$ via Green's theorem [4, Thm. 2.24]: for $\psi, \varphi \in H^2(X)$,

$$(\mathbf{\Omega} \cdot \nabla \psi, \varphi)_{L^2(X)} = (\psi, \varphi)_{L^2(\partial X)} - (\psi, \mathbf{\Omega} \cdot \nabla \varphi)_{L^2(X)},$$

in which case we look in Problem 2 for $\psi \in H^2(X)$ and consider test functions also from $H^2(X)$.

Coercive weak formulations

Under the assumptions that κ is an even function of $\mathbf{\Omega} \cdot \mathbf{\Omega}'$ and q an even function of $\mathbf{\Omega}$, Vladimirov [117] has shown how to obtain the generalized solution of Problem 1 from the solution of

$$\tilde{a}(\tilde{\psi}, \varphi) = \tilde{f}(\varphi) \quad \forall \varphi \in H^p(X) \tag{2.27}$$

with *bounded and coercive* bilinear form $\tilde{a}$ corresponding to (bounded and coercive) transport operator $\tilde{T}$ and source term $\tilde{q}$ of the even parity form of the

NTE[3]. That is, there exist $C_b > 0, \alpha > 0$ such that $\forall u, v \in V$:

$$\tilde{a}(u,v) \leq C_b \|u\|_V \|v\|_V, \quad \tilde{a}(u,u) \geq \alpha \|u\|_V^2$$

and the well-known Lax-Milgram lemma can be used to prove well-posedness of (2.27) and consequently of Problem 1. For the sake of completeness and future reference, we state the lemma below in its general form.

Lemma 1 (Lax-Milgram)**.** *Let V be a Hilbert space and $a : V \times V \to \mathbb{R}$ a bounded and coercive bilinear form with constants C_b and α, respectively. Then for any $f \in V'$, there exists a unique $u \in V$ such that*

$$a(u,v) = f(v) \quad \forall v \in V, \tag{2.28}$$

and

$$\|u\|_V \leq \frac{1}{\alpha} \|f\|_{V'}.$$

Bourhrara presents in [12] three coercive weak formulations of the fixed source problem (and one for the criticality problem investigated in the following section) and shows that one of them actually represents the weak form of another well-known second-order form of the transport equation – the self-adjoint angular flux equation (obtained by expressing ψ from the $\sigma_t \psi$ term of (2.6) in terms of the remaining terms and substituting back into the advection term). He also provides a comparison with other formulations based on the least-squares approach (minimizing the appropriately scaled residual $\|\mathcal{P}(T\psi - q)\|^2_{L^2(X)} + \varpi \|\psi_{\text{in}} - \psi\|^2_{L^2(\partial X^-)}$) studied e.g. in [79] or [4].

Approximations of Problem 2 leading to practical numerical solution methods are naturally obtained by restricting the formulation to finite-dimensional subspaces of $H^2(X)$. With bounded and coercive bilinear form a, approximation error is then automatically assessed by the Céa's lemma (e.g., [107, 2.1.2]). In the angular domain, this has been done using the subspace of spherical harmonics of finite degree in [14] of [79], which is also known as the P_N method (the finite-dimensional restriction can then be obtained for instance by using finite-element

[3]This well-known and widely utilized form of the transport equation is obtained by writing eq. (2.6) for $\mathbf{\Omega}$ and $-\mathbf{\Omega}$, adding and subtracting the two resulting equations and eliminating the unknowns of odd parity; see e.g. [9, Chap. II], [109, Sec. 9.11] or [30, Chap. XX]).

discretization of the spatial dependence). The P_N method will be the subject of Sec. 3.3; as we will see in Sec. 3.4, the same principle also applies to the other most widely employed angular approximation method, the S_N method.

2.2.5.4 Subcriticality conditions

For future reference, we collect the assumptions of Theorem 1 into the following definition.

Definition 1 (Subcriticality conditions)**.** Let

(a) $\sigma_t \in L^\infty(X)$, $\sigma_t \geq \underline{\sigma_t} > 0$ a.e. in $\mathcal{D} \times [E_{\min}, E_{\max}]$,

(b) $\kappa \geq 0$ a.e. in $\mathcal{D} \times \mathcal{S}_2^2 \times [E_{\min}, E_{\max}]^2$,

(c) $c \leq \bar{c} < 1$ a.e. in $\mathcal{D} \times [E_{\min}, E_{\max}]$ where

$$c(\mathbf{r}, E) = \frac{1}{\sigma_t(\mathbf{r}, E)} \int_{E_{\min}}^{E_{\max}} \int_{\mathcal{S}_2} \kappa(\mathbf{r}, \boldsymbol{\Omega}' \cdot \boldsymbol{\Omega}, E' \leftarrow E) \, \mathrm{d}E' \, \mathrm{d}\boldsymbol{\Omega}' \,,$$

(d) $d \leq \overline{d} < 1$ a.e. in $\mathcal{D} \times [E_{\min}, E_{\max}]$ where

$$d(\mathbf{r}, E) = \frac{1}{\sigma_t(\mathbf{r}, E)} \int_{E_{\min}}^{E_{\max}} \int_{\mathcal{S}_2} \kappa(\mathbf{r}, \boldsymbol{\Omega} \cdot \boldsymbol{\Omega}', E \leftarrow E') \, \mathrm{d}E' \, \mathrm{d}\boldsymbol{\Omega}' \,.$$

Then we call

- conditions (a,b,c) the subcriticality conditions in $L^1(X)$,
- conditions (a,b,d) the subcriticality conditions in $L^\infty(X)$,
- conditions (a-d) the subcriticality conditions in $L^p(X)$, $2 \leq p < \infty$.

2.2.6 Criticality problem

The other important problem in neutron transport (particularly in nuclear reactor engineering) requires the determination of material composition (i.e. the values of σ_x) for a given domain geometry (or vice versa) that would ensure a steady neutron distribution (that means – steady power generation) with no additional neutron sources besides fission. This is called a "criticality problem" – the system is said to be *subcritical*, *supercritical* and *critical*, respectively, if without an additional neutron source the number of neutrons in the core will, respectively, continuously diminish, increase or be maintained through the balance between actual out of core leakage, absorption and fission. This characterization motivates the name of the conditions in Def. 1. In reactor core reloading optimization, we assume the core geometry fixed and try to find such a material composition that (besides other optimization criteria) would ensure the critical state.

Mathematically, we are looking for a non-trivial non-negative solution of the homogeneous version of eq. (2.6) (i.e. with $q \equiv 0$ and boundary conditions (2.8)), which means solving an eigenvalue problem. The resulting eigenvalue then describes the departure from critical state with the current set of material data and the associated eigenfunction represents the shape of neutron flux in such a steady state.

In order to formulate the eigenvalue problem, we split the kernel of the collision operator according to (2.12) into the scattering and fission part, thus

$$K\psi = S\psi + F\psi$$

where, using further eq. (2.15),

$$\begin{aligned} F\psi(\mathbf{r},\boldsymbol{\Omega},E) &= \frac{\chi(E)}{4\pi}\int_{E_{\min}}^{E_{\max}} \nu\sigma_f(\mathbf{r},E') \int_{\mathcal{S}_2} \psi(\mathbf{r},\boldsymbol{\Omega},E')\,\mathrm{d}\boldsymbol{\Omega}\;\mathrm{d}E' \\ S\psi(\mathbf{r},\boldsymbol{\Omega},E) &= \int_{E_{\min}}^{E_{\max}} \int_{\mathcal{S}_2} \eta\sigma_s(\mathbf{r},\boldsymbol{\Omega}\cdot\boldsymbol{\Omega}', E\leftarrow E')\psi(\mathbf{r},\boldsymbol{\Omega},E')\,\mathrm{d}\boldsymbol{\Omega}'\;\mathrm{d}E' \end{aligned} \tag{2.29}$$

The criticality eigenvalue problem then reads:

Problem 3. *Find nontrivial, non-negative* $\psi \in \mathrm{Dom}\,(B) \subset L^p(X)$ *and* $\lambda > 0$*, such that*

$$\begin{cases} B\psi \equiv \big[A + \Sigma_t - S\big]\psi = \dfrac{1}{\lambda} F\psi, \\ \mathrm{Dom}\,(B) = \{\psi \in H^p(X),\ \psi|_{\partial X^-} = B_\beta \psi\}, \end{cases} \tag{2.30}$$

where $B_\beta\psi$ *is given by the right-hand side of* (2.8) *(or, more generally,* B_β *is the boundary operator of* (2.9)*).*

Proving existence and uniqueness of solution of (2.30) can proceed in the following sequence:

1. Prove that the transport operator B is invertible. This permits the traditional transcription of the eigenvalue equation (2.30):

   $$B^{-1}F\psi = \lambda\psi.$$

   Results of the previous section can be used here if we consider the collision kernel κ without the fission part, e.g. if instead of the average number of all emitted neutrons in (2.23) we consider only the number emitted from scattering collisions (the *scattering ratio*):

   $$\tilde{c}(\mathbf{r}, E) \coloneqq \frac{1}{\sigma_t(\mathbf{r}, E)} \int_{E_{\min}}^{E_{\max}} \int_{\mathcal{S}_2} \sigma_s(\mathbf{r}, \boldsymbol{\Omega}\cdot\boldsymbol{\Omega}', E' \leftarrow E)\, \mathrm{d}E'\, \mathrm{d}\boldsymbol{\Omega}'\,. \tag{2.31}$$

2. Prove that operator $B^{-1}F$ is (strongly) positive and compact. As such, it has countably many eigenfunctions. Positivity can be deduced from physical properties of the involved operators (although this may place much too severe restrictions on the coefficients, see below). Compactness is harder to establish and holds in general in $L^p(X)$ spaces for $1 < p < \infty$, but not for $p = 1$ or $p = \infty$ (we will comment on this case below).

3. Invoke the Krein-Rutmann theorem for positive linear compact operators (e.g., [39, Thm. 5.4.33]) to prove that the spectral radius of $B^{-1}F$ is a simple eigenvalue associated with the unique positive eigenfunction.

Remark 5 (Criticality). In nuclear engineering, spectral radius of $B^{-1}F$ is often denoted k_{eff} and called *effective multiplication factor*. Physically,

$$k_{\text{eff}} \equiv \frac{\text{neutron emission}}{\text{neutron loss}}$$

so that $k_{\text{eff}} < 1$, $k_{\text{eff}} > 1$ and $k_{\text{eff}} = 1$, respectively, correspond to *subcritical*, *supercritical* and *critical* system. Notice that $c < 1 \Rightarrow k_{\text{eff}} < 1$, but not the other way round because the possibility of neutron loss due to out of core leakage is not accounted for in (2.23).

Because of the first step, we can expect similar assumptions as in Thm. 1 (or in the discussion below the theorem) to be required. Depending on the chosen functional setting, various additional assumptions need to be made in order to carry out the other two steps. These mathematical assumptions restrict either the boundary conditions, geometry or material composition of the solution domain, or energetic dependence (or all) and may not always coincide with physical reality. For instance, strong positivity of $B^{-1}F$ would require $\sigma_f \geq \underline{\sigma_f} > 0$ a.e. in X, implying that fission occurs everywhere, which it generally does not (consider for instance the area between fuel rods in nuclear reactors, filled with coolant water). For only a non-strongly positive compact operator, one can still use the weak form of the Krein-Rutmann theorem ([39, Prop. 5.4.32]). That theorem however does not guarantee uniqueness of the eigensolution and a separate demonstration is required. In $L^1(X)$ or $L^\infty(X)$, compactness of $B^{-1}F$ can be replaced by power compactness, i.e. $(B^{-1}F)^2$ ([101]).

In [101], the above scheme is carried out in the weighted L^1_σ setting introduced in previous section. The result is the following theorem:

Theorem 2. *Let $\tilde{c} < 1$ a.e. in X and either $\sigma_f \geq \tilde{\sigma}_f > 0$ a.e. in X, or at least in a nonempty subset $X_F \subset X$ that is trajectory-connected with whole X (see below). Further assume that S and F can be approximated by compact operators S_n and F_n, respectively:*

$$\lim_{n\to\infty} \|S - S_n\|_{\mathcal{L}(L^1_\sigma, L^1)} = \lim_{n\to\infty} \|F - F_n\|_{\mathcal{L}(L^1_\sigma, L^1)} = 0 \tag{2.32}$$

Then the problem

$$B\psi = \frac{1}{\lambda} F\psi, \quad \text{Dom}\,(B) = \{\psi \in H^1_\sigma(X),\ \psi|_{\partial X^-} = B_\beta \psi\},$$

where $B_\beta : L^1_\sigma(\partial X^+) \to L^1_\sigma(\partial X^-)$ is the albedo operator of (2.9), has a countable number of eigenvalues $\{\lambda_k\}$ and associated (generalized) eigenfunctions which belong to $H^1_\sigma(X)$. There exists the eigenvalue $\lambda = \min\{\lambda_k\} = \rho(B^{-1}F)$ (the spectral

radius of $B^{-1}F$), which is algebraically simple and its associated eigenfunction is the only one that does not change sign in X.

Proof. See [101]. □

The condition $\tilde{c} < 1$ may be violated if η is considerably larger than 1, i.e. in case of high neutron yield from non-elastic scattering; this case is however reasonably excluded by physical reasons (at least in thermal reactor calculations). The second condition is needed for uniqueness – the notion of trajectory connectivity is so far rather heuristic and basically means that particles produced in X_F may reach any other point by direct streaming or through collisions. Essentially similar conditions are often used to circumvent the unphysical restriction of almost everywhere strictly positive fission cross-sections ([5, 87]). Assumption (2.32) is needed for proving power compactness of $B^{-1}F$ and is physically non-restrictive as it requires only uniform continuity of functions that characterize probability of transfer from $(\mathbf{\Omega}, E) \leftarrow (\mathbf{\Omega}', E')$ (for σ_f, this is actually the function $\chi(E)$ of (2.15)) and not that of physical cross-sections $\sigma_{sn}(\mathbf{r}, E')$ and $\sigma_{fn}(\mathbf{r}, E')$ themselves (see [101]).

2.2.7 Rotational invariance of the NTE

An important property of the neutron transport equation is its orthogonal invariance, which says that under certain circumstances, to obtain a solution of the NTE with source term rotated (or reflected) around origin it is sufficient to apply the same rotation (reflection) on the solution corresponding to the original source. We will henceforth consider only rotations but any argument below applies also for reflections.

Definition 2. We will say that an operator equation

$$\mathcal{A}u = f \tag{2.33}$$

is rotationally invariant, if $\mathcal{A}u = f$ implies $\mathcal{A}\mathcal{R}u = \mathcal{R}f$ for any operator $\mathcal{R}$ corresponding to a rotation $\mathbf{R} \in \mathbb{R}^{3\times3}$ of coordinate system around origin[4]:

$$\begin{gathered} \mathcal{R} : f(\mathbf{r}, \mathbf{\Omega}) \mapsto f(\mathbf{R}^T\mathbf{r}, \mathbf{R}^T\mathbf{\Omega}) \\ \mathbf{R}^T\mathbf{R} = \mathbf{R}\mathbf{R}^T = \mathbf{I}, \quad \det \mathbf{R} = 1. \end{gathered} \tag{2.34}$$

[4]by $\mathbf{I}$, we will henceforth denote the unit matrix of appropriate size obvious from the context

Operator $\mathcal{R}$ is defined by its associated rotation matrix $\mathbf{R}$, which is conventionally characterized as an element of the special orthogonal group in $\mathbb{R}^3$, the SO(3) [5]. We will also consider the operator itself to be an element of that group, i.e. $\mathcal{R} \in \mathrm{SO}(3)$. The following lemma shows that equation (2.33) is rotationally invariant if and only if its operator commutes with rotations.

Lemma 2. $\mathcal{A}u = f \;\Rightarrow \mathcal{A}\mathcal{R}u = \mathcal{R}f \;\; \forall \mathcal{R} \in \mathrm{SO}(3)$ *if and only if* $\mathcal{A}\mathcal{R} = \mathcal{R}\mathcal{A}$.

Proof. Sufficiency is obvious by operating with $\mathcal{R}$ on both sides of eq. (2.33). We will show necessity indirectly, i.e. we suppose that there exists $\mathcal{R} \in \mathrm{SO}(3)$ such that $\mathcal{A}\mathcal{R} \neq \mathcal{R}\mathcal{A}$ and show that then we can have $\mathcal{A}u = f$ but $\mathcal{A}\mathcal{R}u \neq \mathcal{R}f$. Indeed, assuming $\mathcal{A}u = f$ and operating by $\mathcal{R}$, we get

$$\mathcal{R}f = \mathcal{R}\mathcal{A}u \neq \mathcal{A}\mathcal{R}u.$$

□

Using definitions from previous subsections, let us write eq. (2.6) in the form of (2.33):

$$T\psi \equiv (L - K)\psi = q, \tag{2.35}$$

where we now suppose generally $T : V \to V$ for some suitable function space V in which we have assured existence of unique solution of (2.35) for $q \in V$ (see Sec. 2.2.5 for examples) . Let us also consider $\mathcal{R}$ as an operator from V into itself. Then the following claim (by Zweifel and Case, [20, Theorem 3]) is valid.

Theorem 3. *If the coefficient functions σ and κ are invariant under the action of $\mathcal{R}$, then also*

$$\mathcal{R}T = T\mathcal{R}. \tag{2.36}$$

Proof. Because A is represented by a dot product of two vectors and Σ_t is rotation invariant as a consequence of the assumptions, L is rotation invariant as well. Commutativity of K and $\mathcal{R}$ follows again from rotational invariance of dot product (i.e., $\mathbf{R}^T\boldsymbol{\Omega} \cdot \boldsymbol{\Omega}' = \boldsymbol{\Omega} \cdot \mathbf{R}\boldsymbol{\Omega}'$), substitution $\mathbf{R}\boldsymbol{\Omega}' = \boldsymbol{\Omega}''$ in the angular integral and the fact that Jacobian determinant of this orthogonal transformation is unity. □

[5] or just the ordinary orthogonal group $O(3)$ if reflections are taken into account

Assumptions of the theorem will be satisfied for instance in an isotropic homogeneous region where $\sigma_t(\mathbf{r}, E) \equiv \sigma_t(E)$ and invariance of κ follows from (2.5). If we also assume that sources (and boundary conditions) are rotationally invariant ($q = \mathcal{R}q$), then the true solution of NTE is (by Lemma 2) necessarily spherically symmetric or, in other words, the rotated solution satisfies the original equation:

$$T\psi = q \quad \Rightarrow \quad T\mathcal{R}\psi = q.$$

Numerical approximations should preserve this property in order to produce physically correct results. As we will see in the following chapter, however, this is not the case for the widely used S_N approximation and leads to undesirable numerical side-effects.

3

Neutron transport approximations

As mentioned in the introductory chapter, we will focus on deterministic methods for solving the NTE, requiring proper discretization of (2.6) and solution of the resulting system of algebraic equations. In the following sections, we review some of the most widely used semi-discretizations with respect to energy, angle and spatial variables and put them into a unified Hilbert space projection framework. We finish this chapter with a general discussion on solving the associated large sparse systems of algebraic equations We will mostly concern ourselves with the fixed-source problem; solving this problem is, however, a necessary part of practically all numerical methods for solving the generalized eigenvalue problem (2.30) as well.

Notation conventions

Concerning fonts and subscripts/superscripts, we will generally use the following conventions (wherever exception will be needed, it will be clearly stated):

- $\mathbf{f}$ ... column vector with numerical values ($\mathbf{f} \in \mathbb{R}^N$ for some $N \in \mathbb{N}$);
- f_n or $[\mathbf{f}]_n$ ... components of $\mathbf{f}$;
- $\mathbf{A}$ ... matrix with numerical values ($\mathbf{A} \in \mathbb{R}^{M\times N}$, $M, N \in \mathbb{N}$); also, $\mathbf{A}(\mathbf{r})$ will denote matrix-valued function with elements being functions in $\mathcal{D}$;

- A_{ij} or $[\mathbf{A}]_{ij}$ ... elements of $\mathbf{A}$;
- upshape F (including Ψ, Φ) ... vector-valued function $\mathcal{D} \to \mathbb{R}^N$, $N \in \mathbb{N}$; as an exception, $\mathbf{n}(\mathbf{r})$ denotes as before the vector-valued function representing unit outward normal field of $\partial\mathcal{D}$;
- f_n or $[\mathrm{F}]_n$ ... components of F;
- A or calligraphic $\mathcal{A}$ ... in the context of an operator acting on some vector space V, usual letters will be used for transport operators introduced in previous chapter, while calligraphic letters for general operators;
- $s = \{c_k\}_{k=1}^N \equiv \{c_k\}_N$;
- col s ... column vector with entries $c_1, c_2, \ldots, c_N$;
- diag s ... diagonal matrix whose diagonal is given by elements of s;
- $f_{(i)}$ ... i-th iterate in an iteration process.

So with this notation, we have, for instance, $\mathrm{F} = \operatorname{col}\{f_n\}_N$ with each f_n being a function from some function space $V_n(\mathcal{D})$.

To facilitate comparison of the results with literature, we also neglect inelastic scattering (i.e., put $\eta \equiv 1$). The scattering component of the collision kernel (first relation in (2.13)) then becomes

$$\sigma_s(\mathbf{r}, E) = \int_{E_{\min}}^{E_{\max}} \int_{\mathcal{S}_2} \sigma_s(\mathbf{r}, \boldsymbol{\Omega}' \cdot \boldsymbol{\Omega}, E' \leftarrow E)\, \mathrm{d}\boldsymbol{\Omega}'\, \mathrm{d}E' \,. \tag{3.1}$$

3.1 Approximation of energetic dependence

The continuous dependence on energy, $\psi = \psi(\cdot, \cdot, E)$, is typically resolved by the so called *multigroup approximation*. In this approximation, the interval of neutron energies is divided as follows:

$$\begin{aligned}\left[E_{\min}, E_{\max}\right] &= \left[E^G - \tfrac{\Delta E^G}{2}, E^G + \tfrac{\Delta E^G}{2}\right] \cup \ldots \\ &\ldots \cup \left[E^g - \tfrac{\Delta E^g}{2}, E^g + \tfrac{\Delta E^g}{2}\right] \cup \ldots \cup \left[E^2 - \tfrac{\Delta E^2}{2}, E^2 + \tfrac{\Delta E^2}{2}\right] \cup \left[E^1 - \tfrac{\Delta E^1}{2}, E^1 + \tfrac{\Delta E^1}{2}\right],\end{aligned}$$

where $E^{g+1}+\frac{\Delta E^{g+1}}{2}=E^g-\frac{\Delta E^g}{2}$, and equations (2.6–2.8) are integrated over each energy group range $\left[E^g-\frac{\Delta E^g}{2},E^g+\frac{\Delta E^g}{2}\right]$.

REMARK 6. Note that the energy intervals (groups) are traditionally numbered in a descending order, i.e. a group with larger index contains lower energies than a group with lesser index; also, the group index is traditionally placed in superscript.

The NTE (2.6) is thus transformed into a finite system of integro-differential equations, each governing the flux of neutrons with energies within a particular range (in this context called *group*):

$$\psi^g(\mathbf{r},\mathbf{\Omega})=\frac{1}{\Delta E^g}\int_g\psi(\mathbf{r},\mathbf{\Omega},E),\mathrm{d}E\equiv\frac{1}{\Delta E^g}\int_{E^g-\Delta E^g/2}^{E^g+\Delta E^g/2}\psi(\mathbf{r},\mathbf{\Omega},E),\mathrm{d}E\,,\qquad g=1,2,\dots G. \tag{3.2}$$

This conventional procedure leads to the following set of G coupled neutron transport equations

$$\begin{cases} T_G\{\psi^g\}_G=\{q^g\}_G,\\ \mathrm{Dom}\,(T_G)=\left\{\{\psi^g\}_G\in\left[H^p(X_{|E})\right]^G,\ \psi^g|_{\partial X^-_{|E}}=0,\ g=1,\dots,G\right\},\end{cases} \tag{3.3}$$

where

$$X_{|E}\coloneqq\{(\mathbf{r},\mathbf{\Omega}):\ \mathbf{r}\in\mathcal{D}\subset\mathbb{R}^3,\mathbf{\Omega}\in\mathcal{S}_2\}$$

is the 5-dimensional subspace of X (i.e., the norm in $H^p(X_{|E})$ is defined as in (2.21) but only using double integrals over $\mathcal{D}\times\mathcal{S}_2$) and analogously for $\partial X^{\pm}_{|E}$. The multigroup transport operator has the following form:

$$T_G\{\psi^g\}_G=\left\{(A+\Sigma_r^g)\,\psi^g-\sum_{g'=1,g'\neq g}^{G}K^{gg'}\psi^{g'}\right\}_G,$$

$$\Sigma_r^g\psi^g(\mathbf{r},\mathbf{\Omega})=\sigma_t^g(\mathbf{r})\psi^g(\mathbf{r},\mathbf{\Omega})-\int_{\mathcal{S}_2}\kappa^{g\leftarrow g}(\mathbf{r},\mathbf{\Omega}\cdot\mathbf{\Omega}')\psi^g(\mathbf{r},\mathbf{\Omega}')\,\mathrm{d}\mathbf{\Omega}',$$

$$K^{gg'}\psi^{g'}(\mathbf{r},\mathbf{\Omega})=\int_{\mathcal{S}_2}\kappa^{g\leftarrow g'}(\mathbf{r},\mathbf{\Omega}\cdot\mathbf{\Omega}')\psi^{g'}(\mathbf{r},\mathbf{\Omega}')\,\mathrm{d}\mathbf{\Omega}'$$

where the terms with superscript g or g' represent quantities suitably averaged over $\left[E^g-\frac{\Delta E^g}{2},E^g+\frac{\Delta E^g}{2}\right]$, e.g. $k^{g\leftarrow g'}$ is (in theory) obtained as

$$\kappa^{g\leftarrow g'}(\mathbf{r},\mathbf{\Omega}\cdot\mathbf{\Omega}')=\frac{\int_g\int_{g'}\kappa(\mathbf{r},\mathbf{\Omega}\cdot\mathbf{\Omega}',E\leftarrow E')\psi(\mathbf{r},\mathbf{\Omega},E')\,\mathrm{d}E'\,\mathrm{d}E}{\int_g\psi(\mathbf{r},\mathbf{\Omega}',E)\,\mathrm{d}E}. \tag{3.4}$$

It is customary to move the *self-scattering* (diagonal) part of the collision operator to the reaction operator. Since the reactions in which energetic distribution of both the incoming and outgoing neutrons lies within the same group are included in both σ_t and κ (compare equations (2.13) and (2.14)), this transformation makes $\Sigma_r^g \psi^g$ represent the actual rate of neutron removal from the group, while $K^{gg'}\psi^{g'}$ the rate of neutron addition to that group. Results about unique solvability presented in previous chapter carry over to the multigroup setting by considering a counting measure on the set $\{E^G, \ldots, E^1\}$ instead of the continuous Lebesgue measure $\mathrm{d}E$ [30, Chap. XXI §2].

3.1.1 Multigroup data

Although the multigroup system of neutron transport equations has a relatively simple form, finding an optimal grouping of energies and determining the associated group-averaged coefficients is not an easy task in most practical applications because of the highly complicated energetic dependence of nuclear processes, as illustrated by the typical dependence of the (microscopic) fission cross-section of ${}^{235}_{92}\mathrm{U}$ in the so-called resonance range of energies and corresponding multigroup approximation in Fig. 3.1. Suitable approximation of the unknown exact solution in (3.4) is also highly non-trivial, albeit essential for the success of the multigroup method. Even though an alternative to the finite-volume like approximation (3.2) has been proposed recently in [37] – using Galerkin projection of angular flux onto a space of functions supported over subregions of the energy range (a finite-element like approach) – the multigroup approximation still remains the most universally used approach to simplify the energetic dependence (see, e.g., [18, Chap. 5] or [24]). However, we will not specifically address this issue and always assume that the multigroup coefficients appearing in the equations are given as input.

Remark 7. Fission spectrum In criticality problems, the set of multigroup data must include both parts of the collision kernel $\kappa^{gg'}$, i.e. the cross-sections $\sigma_s^{gg'}$ and $\sigma_f^{gg'}$, as well as $\nu^{g'}$ and χ^g. Because of the rapid decay of $\chi(E)$ for low energies (as neutrons are mostly emitted from fission with high energies) that are nevertheless determining for the cross-sections (as most interactions are likely to

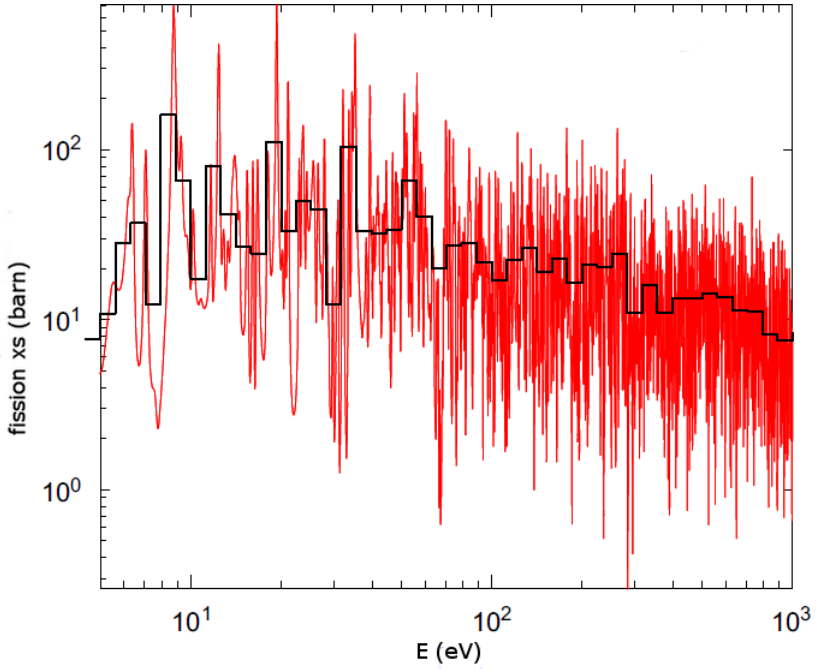


Figure 3.1: Microscopic fission cross-section of U235.

occur due to slowly moving neutrons, at least in classical moderated reactors)[1], there will typically be $\chi^g = 0$ for $g = G, G-1, \ldots, G-k$ with $k < G$. The group-discretized operator F from (2.29) will therefore have a non-trivial null-space, leading ultimately to a fully discrete partial generalized eigenproblem

$$\text{Find } (\lambda_{\min}, \mathbf{x}) \text{ where } \lambda_{\min} \text{ is minimal } \lambda \in \mathbb{R}^+ \text{ such that } \mathbf{A}\mathbf{x} = \lambda \mathbf{B}\mathbf{x}, \quad \mathbf{x} \neq \mathbf{o} \tag{3.5}$$

with singular $\mathbf{B}$ (which may be solved by the classical shift-and-invert method as described e.g. in [57] or by transformation to the classical eigenvalue problem $\mu\mathbf{x} = \mathbf{A}^{-1}\mathbf{B}\mathbf{x}$ for the dominant eigenvalue $\mu = 1/\lambda_{\min} = k_{\text{eff}}$).

3.1.2 Group source iteration

A standard way of iterative solution of the multigroup system is the *group source iteration*:

[1] cf. Fig. 3.1 and Fig. 2.6 and notice the different scaling on the horizontal axis

For a given initial approximation $\psi^g_{(0)}$, $g = 1, \ldots, G$, solve

for i = 0,1,...

for g = 1,...,G

$$(A + \Sigma^g_r)\,\psi^g_{(i+1)} = \sum_{g' \le g-1} K^{gg'}\psi^{g'}_{(i+1)} + \sum_{g' \ge g+1} K^{gg'}\psi^{g'}_{(i)} + q^g. \qquad (3.6)$$

If we view the operator T_G as a matrix operator acting on $\mathrm{col}\,\{\psi^g\}_G$, then we can interpret this iteration as a Gauss-Seidel iteration for (3.3), where T_G has been split into its lower-triangular part $A + \Sigma^g_r - K^{gg'}$ $(g' \le g)$ and its upper triangular part $K^{gg'}$ $(g' > g)$ and the lower triangular part is being inverted by forward substitution. Convergence of this scheme can become slow when the upper triangular part (representing neutron up-scattering from lower energies to higher energies) is dominating. Therefore, when preparing the multigroup data, it is advantageous to put an effort into finding such an energy grouping that minimizes the up-scattering (which is often done in practice, as in e.g. [69]).

Remark 8. Here we assume that the mono-energetic problem can be solved exactly. Approximations of angular dependence discussed in the following section (like S_N) may employ another iteration level to resolve angular coupling of the within-group fluxes induced by collisions. This iteration is usually called just *source iteration* and can also become slow if scattering of neutrons with given energy dominates their absorption (we will return to this issue later in Sec. 3.4.3).

Note that by employing the group source iteration, only a mono-energetic transport problem in group g has to be solved in each iteration, and if the differential part A can be represented by a symmetric operator $\tilde{A}$ (as can be done for some of the second-order forms described in Sec. 2.2.5.3 or when suitable angular approximations like diffusion are being used – see Sec. 3.3.5), the problem would become symmetric with implications for efficient numerical solution. In the remainder of this chapter, we will focus on the approximation of neutron flux in a single group (index of which will be omitted), described by the corresponding within-group equation in which contributions from other groups are encapsulated in the source term q (i.e., we will study the NTE on $X_{|E}$). In order to simplify notation, we will use just X instead of $X_{|E}$ when referring to the solution domain.

3.2 Approximation of angular dependence

Considerably larger number of methods have been proposed for approximating the angular dependence of neutron flux. Many of them are still being used and actively developed today as their characteristics make them more suitable for one application area than other methods, which are preferred in different areas.

3.2.1 Methods based on the integral form of NTE

As a first example, we consider the class of methods originally derived from the equivalent integral form of the NTE (see Sec. 2.2.1). Typical representatives of this class are the method of collision probabilities or the method of characteristics (see e.g. [23, 62, 104, 122]; the computer code DRAGON [80] used as a reference transport solver in Sec. 6.3 is also based on these two methods). As the integral form of the NTE represents global neutron balance over the domain, the corresponding algebraic systems (obtained after spatial discretization) are full and their solution demanding on computer resources. On the other hand, these methods quite naturally handle complex geometries and are well suited for smaller-scale, high-fidelity calculations[2] indispensable for generating appropriately averaged coefficients for the larger scale (whole-core) calculations. This *spatial homogenization* and *energy group condensation*, as these averaging procedures are traditionally called in nuclear engineering field, are employed by many existing whole-core simulators (see e.g. [97, Chap. 17] or the review in first two sections of [102]). To simulate long-term nuclear reactor operation, it is furthermore necessary to perform these procedures under varying physical conditions of the core and generate many sets of averaged coefficients corresponding to these conditions.

[2] called *lattice calculations* as they are typically performed on a single representative subdomain of the core (one or several neighboring assemblies of fuel elements (pins), or the fuel element itself) with reflective boundary conditions, simulating an infinite lattice of such subdomains

3.2.2 Methods based on the integro-differential NTE

More suitable for whole-core calculations are methods derived from the integro-differential version of the NTE, eq. (2.6), that lead to sparse algebraic systems. The best-established are the method of discrete ordinates (S_N) and the method of spherical harmonics (P_N). Both arise from applying in the angular domain a classical well known approach for constructing finite numerical approximations of PDEs. In the following sections, we will briefly introduce the main ideas behind the S_N and P_N methods and expose their most important properties. These properties are generally known, but their origin in mathematical structure of the approximate forms is often overlooked in literature (a few exceptions will be cited below and in the corresponding appendices).

We will also interpret both methods as restrictions of the same continuous NTE onto appropriate closed semi-finite dimensional subspaces of $H^2(X)$. This is automatic in the case of the P_N approximation (which is basically a Galerkin method in angular domain with globally supported continuous basis functions), but has (as far as the author knows) not been explicitly done for the S_N approximation. This will be the subject of Sec. 3.4.2 for the practically important case of isotropic scattering, i.e. $\sigma_s(\mathbf{r}, \mathbf{\Omega} \cdot \mathbf{\Omega}') = \frac{\sigma_s(\mathbf{r})}{4\pi}$.[3] Apart from showing both methods in the same light, this approach also allows to use properties of the continuous transport operators to analyze the behavior of the approximate methods, as will be illustrated in Sec. 3.4.3.

3.3 The P_N method

The system of P_N equations has been originally derived using the weighted residuals method in the angular domain. That is, the angular flux is expanded into infinite series of functions of $\mathbf{\Omega}$ that span a complete basis on the unit sphere, the continuous neutron transport equation (2.6) is multiplied by each member of the basis in turn and integrated over the sphere. The properties of the basis functions are then used to derive equations for the expansion coefficients.

[3] The question whether the S_N approximation with general scattering could be rigorously cast as a restriction of the NTE to a subspace of $H^p(X)$ is left open for future investigation.

Only a finite number of expansion terms is considered to allow practical computation. Usually, the expansion is truncated to a finite length of $K = K(N)$ terms[4] by setting all expansion coefficients with higher index to 0 (although there exist alternative closure methods that may have favorable properties in certain situations, see e.g. [47]). Then we speak of the P_N approximation:

$$\psi(\mathbf{r}, \mathbf{\Omega}) \approx \sum_{k=1}^{K} \phi_k(\mathbf{r}) \varphi_k(\mathbf{\Omega}). \tag{3.7}$$

A natural function space to support this procedure is the Hilbert space of square-integrable functions on the sphere $L^2(\mathcal{S}_2)$, equipped with the inner product

$$(u, v)_{L^2(\mathcal{S}_2)} = \int_{\mathcal{S}_2} u(\mathbf{\Omega}) v(\mathbf{\Omega}) \, \mathrm{d}\mathbf{\Omega} \, . \tag{3.8}$$

We will therefore assume $\psi(\mathbf{r}, \cdot) \in L^2(\mathcal{S}_2)$ (as is the case, e.g., when $\psi \in H^2(X)$).

The system of spherical basis functions $\{\varphi_k\}_K$ that were used in the original P_N method is the system of *spherical harmonic functions* (shortly *spherical harmonics*, App. A). We will consider here the real system (whose elements are sometimes called *tesseral spherical harmonics*) as it is more convenient for practical purposes than the equivalent complex system (which is more widespread in nuclear engineering literature, e.g. [109, Sec. 9.7], [97, Sec. 14.4]), [110, Chap. V]).

In one dimension, spherical harmonics reduce to Legendre polynomials (A.3) and $K(N) = N$. For general three-dimensional problems, there are $2n+1$ linearly independent spherical harmonics for each degree n and

$$K(N) = \sum_{n=0}^{N} 2n + 1 = (N+1)^2.$$

The approximation (3.7) is usually rewritten as a double sum

$$\psi(\mathbf{r}, \mathbf{\Omega}) \approx \sum_{k=1}^{K} \phi_k(\mathbf{r}) Y_k(\mathbf{\Omega}) \equiv \sum_{n=0}^{N} \sum_{m=-n}^{n} \phi_n^m(\mathbf{r}) Y_n^m(\mathbf{\Omega}) \tag{3.9}$$

[4]The length of the expansion K should not be confused with the operator K introduced in the previous section; it will be always clear from context which meaning the letter K currently has.

where $Y_n^m(\mathbf{\Omega})$ is the spherical harmonic function of degree n and order m and in the first term on right, we consider the single index k ($1 \leq k \leq K$) that covers all the combinations of n and m ($0 \leq n \leq N$, $-n \leq m \leq n$) appearing in the second term.

Spherical harmonics form a complete orthonormal system on $L^2(\mathcal{S}_2)$ with respect to the inner product (3.8) (or its Hermitian variant when complex spherical harmonics are used) and simplify the algebraic manipulations needed to arrive at the relations determining the coefficients ϕ_k (called *angular moments*). These relations comprise a system of K partial differential equations in spatial domain of the following form[5]:

$$\mathbf{A}^x_{P_N} \frac{\partial \Phi(\mathbf{r})}{\partial x} + \mathbf{A}^y_{P_N} \frac{\partial \Phi(\mathbf{r})}{\partial y} + \mathbf{A}^z_{P_N} \frac{\partial \Phi(\mathbf{r})}{\partial z} + \big[\sigma_t(\mathbf{r})\mathbf{I} - \mathbf{K}_{P_N}(\mathbf{r})\big]\Phi(\mathbf{r}) = \mathrm{Q}_{P_N}(\mathbf{r}), \quad (3.10)$$

where

$$\Phi(\mathbf{r}) = \mathrm{col}\,\{\phi_k(\mathbf{r})\}_K \quad \text{and} \quad \mathrm{Q}_{P_N}(\mathbf{r}) = \mathrm{col}\,\{q_k(\mathbf{r})\}_K \quad (3.11)$$

are, respectively, the vector functions of angular flux moments and angular source moments. The Galerkin procedure results in their special form

$$\phi_k(\mathbf{r}) = \int_{\mathcal{S}_2} \psi(\mathbf{r}, \mathbf{\Omega}) Y_k(\mathbf{\Omega})\, \mathrm{d}\mathbf{\Omega}\,, \quad q_k(\mathbf{r}) = \int_{\mathcal{S}_2} q(\mathbf{r}, \mathbf{\Omega}) Y_k(\mathbf{\Omega})\, \mathrm{d}\mathbf{\Omega}\,, \quad (3.12)$$

Remark 9 (Suppression of spatial dependence).
In order to simplify the notation we shall, until Sec. 3.3.4 and when not explicitly stated otherwise, consider all functions and operators with spatial dependence at an arbitrary fixed point $\mathbf{r} \in \mathcal{D}$. This allows us to write e.g. $\psi \in L^2(\mathcal{S}_2)$, $\mathbf{K}_{P_N}$ becomes an ordinary matrix in $\mathbb{R}^{K\times K}$, expressions (3.12) could be rewritten as $\phi_k = \big(\psi, Y_k\big)_{L^2(\mathcal{S}_2)}$ and $q_k = \big(q, Y_k\big)_{L^2(\mathcal{S}_2)}$, respectively, etc.

Using again the double index ($k = {}^m_n$) and the form of spherical harmonics with $n = 0, 1$, we obtain direct correspondence of the first four moments and the physically important quantities defined in Sec. 2.2.3:

$$\phi = \sqrt{4\pi}\phi_0^0, \quad \mathrm{J} = \sqrt{\frac{4\pi}{3}} \begin{bmatrix} \phi_1^1 \\ \phi_1^{-1} \\ \phi_1^0 \end{bmatrix}.$$

[5]Differentiation and integration of vector functions (such as the term $\frac{\partial \Phi(\mathbf{r})}{\partial x}$ in eq. (3.10)) is conventionally understood component-wise.

In view of (3.9) and the completeness and orthogonality properties of spherical harmonics, (3.12) also shows that the angular flux (as a function of $\mathbf{\Omega}$) in the P_N method is actually approximated by its orthogonal projection onto the finite-dimensional subspace $L^2_K(\mathcal{S}_2) \subset L^2(\mathcal{S}_2)$:

$$\psi \approx \Pi_{P_N}\psi, \quad \big(\Pi_{P_N}\psi\big)(\mathbf{\Omega}) \coloneqq \sum_{k=1}^{K} \big(\psi, Y_k\big)_{L^2(\mathcal{S}_2)} Y_k(\mathbf{\Omega}). \tag{3.13}$$

3.3.1 Operator form

Putting the P_N system (3.10) into a form involving the continuous transport operators from eq. (2.35) is now particularly simple. Using (3.13), let us define two mappings that take a vector $\mathrm{F} = \mathrm{col}\,\{f_k\}_K$ to a function $u \in L^2_K(\mathcal{S}_2)$ and vice versa:

$$\big(\mathcal{I}_{P_N}\mathrm{F}\big)(\mathbf{\Omega}) \coloneqq \sum_{k=1}^{K} f_k Y_k(\mathbf{\Omega}), \quad \widehat{\mathcal{I}}_{P_N} u = \mathrm{col}\,\left\{\big(u, Y_k\big)_{L^2(\mathcal{S}_2)}\right\}_K, \tag{3.14}$$

so that, using (3.11) and (3.12),

$$\mathcal{I}_{P_N}\Phi = \mathcal{I}_{P_N}\widehat{\mathcal{I}}_{P_N}\psi = \Pi_{P_N}\psi.$$

The sought form of the P_N system is then

$$\widehat{\mathcal{I}}_{P_N}(L-K)\mathcal{I}_{P_N}\Phi = \mathrm{Q}_{P_N} = \widehat{\mathcal{I}}_{P_N} q \tag{3.15}$$

or, in the angularly continuous domain,

$$\Pi_{P_N}(L-K)\Pi_{P_N}\psi = \Pi_{P_N} q. \tag{3.16}$$

Viewing equation (3.16) as an operator equation in the dual space of $L^2(\mathcal{S}_2)$ (coinciding with $L^2(\mathcal{S}_2)$ by the Riesz representation theorem) and using symmetry of Π_{P_N}, the corresponding weak formulation reads (including again the full spatial dependence)

$$\big((L-K)\Pi_{P_N}\psi, \Pi_{P_N}\varphi\big)_{L^2(X)} = \big(q, \Pi_{P_N}\varphi\big)_{L^2(X)}, \quad \forall \varphi \in L^2(X). \tag{3.17}$$

We have thus obtained the P_N approximate problem as a restriction of Problem 2 to the (closed) subspace $\mathrm{Range}\,\Pi_{P_N}$. Note that this is still an infinite-dimensional problem, as

$$\dim \mathrm{Range}\,\Pi_{P_N} = \dim \mathrm{Span}\,\{Y_k\}_K \times \dim L^2(\mathcal{D}).$$

3.3.2 Structure of the P_N system

3.3.2.1 Advection part

Each of the advection matrices:

$$\left[\mathbf{A}^s_{P_N}\right]_{k,l} = \int_{\mathcal{S}_2} \Omega_s Y_k(\mathbf{\Omega}) Y_l(\mathbf{\Omega}) \,\mathrm{d}\mathbf{\Omega}\,, \quad s \in \{x, y, z\},\ 1 \leq k, l \leq K \tag{3.18}$$

is symmetric and hence for any $\mathbf{n} = [n_x, n_y, n_z]^T$,

$$\mathbf{A}^{\mathbf{n}}_{P_N} = n_x \mathbf{A}^x_{P_N} + n_y \mathbf{A}^y_{P_N} + n_z \mathbf{A}^z_{P_N}$$

is symmetric and diagonalizable with real eigenvalues. The P_N system is thus (strongly) hyperbolic in the sense of [75, Def. 18.1]. The eigenvalues depend on the vector $\mathbf{n}$ only through its length $\|\mathbf{n}\|$ (see Sec. B for an example when $N = 3$), which shows that the P_N system describes radiation propagation at the same speed[6] in any direction. This hints that rotational invariance of the NTE is preserved by the P_N system, as we will directly show in Sec. 3.3.3.

3.3.2.2 Boundary conditions

The eigenstructure of the advection matrices also provides a clue on how many boundary conditions should be prescribed for the P_N system, which is not immediately clear because of the plane wave coupling. Matrix $A^{\mathbf{n}}_{P_N}$ for given N has

- $N(N+1)/2$ positive eigenvalues,
- $N(N+1)/2$ negative eigenvalues and
- $K(N) - N(N+1) = N + 1$ zero eigenvalues,

irrespective of $\mathbf{n}$. If we take $\mathbf{n}$ to be the unit outward normal to $\partial\mathcal{D}$, these eigenvalues correspond, respectively, to outgoing, incoming and tangential neutron radiation waves. In order for the hyperbolic system to be well-posed, we are

[6] Here we consider the P_N system as a steady-state limit of the time-dependent equation, as explained in Sec. B.

allowed to prescribe only the incoming waves, hence we are allowed to specify $N(N+1)/2$ boundary conditions at any point of the boundary.

It is more difficult to determine what the conditions actually look like as the waves generally contain components of all moments ϕ_n^m. On the grounds of physical reasoning (e.g. [99]), variational analysis (e.g. [31]) or most recently ([103]) the equivalence between hyperbolic and elliptic forms of the P_N equations (arising from the second-order forms of the NTE introduced in 2.2.5.3), the agreed upon form of P_N boundary conditions consistent with the present Galerkin framework is obtained (e.g. for the incoming condition (2.7) and again with general spatial dependence):

$$\begin{gathered}\left(\psi_{\mathrm{in}} - \Pi_{P_N}\psi|_{\partial X^-}, Y_p^m\right)_{L^2(\partial X^-)} = 0,\\ p = \begin{Bmatrix} 0,2,4,\ldots,N-1 & \text{for } N \text{ odd} \\ 1,3,5,\ldots,N-1 & \text{for } N \text{ even} \end{Bmatrix},\ -p \le m \le p;\end{gathered} \tag{3.19}$$

that is, as (oblique) projection of the specified incoming angular flux onto $L^2_K(\partial X^-)$, orthogonal to the subspace of $L^2(\partial X^-)$ spanned by spherical harmonics with even/odd degrees, with respect to the inner product

$$(u,v)_{L^2(\partial X^-)} = \int_{\partial\mathcal{D}}\int_{\mathbf{\Omega}\cdot\mathbf{n}<0} |\mathbf{\Omega}\cdot\mathbf{n}|\, u(\mathbf{r},\mathbf{\Omega})v(\mathbf{r},\mathbf{\Omega})\,\mathrm{d}S\,\mathrm{d}\mathbf{\Omega}\,.$$

We will call boundary conditions of this form (as in [31]) *Marshak boundary conditions*[7].

3.3.2.3 Collision part

As we have seen in previous paragraphs, the P_N system (3.10) couples the advected angular moments in the sum involving the advection matrices (although we note that no more than 7 moments are coupled; see B for an example for $N = 3$ or [103, App. A] for general treatment). On the other hand, the collision matrix $\mathbf{K}_{P_N}$ is diagonal as a consequence of the following lemma.

[7] We only make a remark that there is another form of approximate boundary conditions known as the Mark conditions. The relative merit of one over the other is not completely resolved so both are used in practice. We prefer the former as they are consistent with the Galerkin interpretation of the P_N equations.

Lemma 3. *The spherical harmonic functions Y_n^m diagonalize the collision operator K and*

$$KY_n^m = \kappa_n Y_n^m, \quad n = 0, 1, \ldots, \quad -n \leq m \leq n,$$

where

$$\kappa_n = 2\pi \int_{-1}^{1} \kappa(\mu_0) P_n(\mu_0)\, d\mu_0\,, \quad \mu_0 = \boldsymbol{\Omega} \cdot \boldsymbol{\Omega}',$$

is the n-th Legendre moment of the collision kernel κ.[8]

Proof. As we assume the collision kernel to be a square integrable function of the collision cosine $\mu_0 \equiv \cos\vartheta_0 = \boldsymbol{\Omega} \cdot \boldsymbol{\Omega}'$ (see Fig. A.1) and the Legendre polynomials (A.3) form a complete orthogonal system on $L^2([-1,1])$, we can express the collision kernel as a sum of Fourier series

$$\kappa(\mu_0) = \sum_{n=0}^{\infty} \frac{2n+1}{4\pi} \kappa_n P_n(\mu_0), \quad \kappa_n = 2\pi \int_{-1}^{1} \kappa(\mu_0) P_n(\mu_0) \mathrm{d}\mu_0\,. \tag{3.20}$$

Then for any $n = 0, 1, \ldots,\ -n \leq m \leq n$,

$$\begin{aligned} \big(KY_n^m\big)(\boldsymbol{\Omega}) &= \int_{\mathcal{S}_2} \kappa(\boldsymbol{\Omega} \cdot \boldsymbol{\Omega}') Y_n^m(\boldsymbol{\Omega}')\, \mathrm{d}\boldsymbol{\Omega}' \\ &= \int_{\mathcal{S}_2} \sum_{p=0}^{\infty} \frac{2p+1}{4\pi} \kappa_p P_p(\boldsymbol{\Omega} \cdot \boldsymbol{\Omega}') Y_n^m(\boldsymbol{\Omega}')\, \mathrm{d}\boldsymbol{\Omega}' \\ &= \sum_{p=0}^{\infty} \kappa_p \sum_{q=-p}^{p} Y_p^q(\boldsymbol{\Omega}) \int_{\mathcal{S}_2} Y_p^q(\boldsymbol{\Omega}') Y_n^m(\boldsymbol{\Omega}')\, \mathrm{d}\boldsymbol{\Omega}' \\ &= \kappa_n Y_n^m(\boldsymbol{\Omega}). \end{aligned} \tag{3.21}$$

where the addition theorem (A.5) has been used on third line and orthogonality relation (A.4) on the fourth, completing the proof. □

Corollary 1. *Matrix $\mathbf{K}_{P_N} = \widehat{\mathcal{I}}_{P_N} K \mathcal{I}_{P_N}$ is diagonal, with entries given by the (repeated) Legendre moments κ_n.*

Proof. The j-th column of $\mathbf{K}_{P_N}$ is given by

$$\mathbf{K}_{P_N} \mathbf{e}_j = \widehat{\mathcal{I}}_{P_N} K \mathcal{I}_{P_N} \mathbf{e}_j,$$

[8] we keep in mind that spatial dependence of κ and κ_n is not explicitly shown

where $\mathbf{e}_j$ is the j-th canonical basis vector in $\mathbb{R}^K$. By definition, each index $j = 0, 1, \dots K$ corresponds to a unique double index $\begin{smallmatrix} m \\ n \end{smallmatrix}$ ($n = 0, 1, \dots N$, $-n \leq m \leq n$), so that $Y_j \equiv Y_n^m$. We can therefore write

$$\left(\mathcal{I}_{P_N}\mathbf{e}_j\right)(\mathbf{\Omega}) = Y_j(\mathbf{\Omega}) = Y_n^m(\mathbf{\Omega}).$$

Using Lemma 3, we have

$$K\mathcal{I}_{P_N}\mathbf{e}_j = \kappa_n Y_n^m$$

so, when associating the index i to a double index $\begin{smallmatrix} s \\ r \end{smallmatrix}$ and using the orthogonality relation (A.4),

$$[\mathbf{K}_{P_N}]_{ij} = \left[\widehat{\mathcal{I}}_{P_N} K \mathcal{I}_{P_N}\mathbf{e}_j\right]_i = (\kappa_n Y_n^m, Y_r^s)_{L^2(\mathcal{S}_2)} = \kappa_n \delta_n^r \delta_m^s = \kappa_n \delta_{ij}.$$

Note that there is a single value κ_n for all $2n+1$ functions Y_n^m, so $\mathbf{K}_{P_N}$ consists of $N+1$ diagonal blocks with elements κ_0, 3 times κ_1, 5 times κ_3, etc. □

Corollary 2. *The complete "capture" matrix*

$$\mathbf{C}_{P_N} = \widehat{\mathcal{I}}_{P_N} C \mathcal{I}_{P_N} \equiv \widehat{\mathcal{I}}_{P_N}(\Sigma_t - K)\mathcal{I}_{P_N}$$

(corresponding to the capture cross-section σ_c in (2.14) *and characterizing net neutron loss due to all types of neutron-nuclei interactions) is diagonal.*

3.3.2.4 Legendre scattering moments

For any arbitrary incoming direction $\mathbf{\Omega} \in \mathcal{S}_2$, the 0-th Legendre moment of the scattering component of the collision kernel κ (eq. (3.1)) is

$$\begin{aligned}\sigma_{s0} &= 2\pi \int_{-1}^{1} \sigma_s(\mu_0) P_0(\mu_0) \mathrm{d}\mu_0 \\ &= \int_0^{2\pi}\int_0^{\pi} \sigma_s(\cos\vartheta_0)\sin\vartheta_0 \mathrm{d}\vartheta_0 = \int_{\mathcal{S}_2} \sigma_s(\mathbf{\Omega}'\cdot\mathbf{\Omega})\mathrm{d}\mathbf{\Omega}'.\end{aligned}$$

(where the rule (2.4) has been used). Comparing with def. (3.1),

$$\sigma_{s0} = \sigma_s,$$

i.e., the 0-th Legendre moment of scattering is just the ordinary scattering cross-section. It can be also shown that

$$\sigma_{s1} = \bar{\mu}_0 \sigma_s,$$

where $\bar{\mu}_0$ is the mean scattering cosine. σ_s and $\bar{\mu}_0$ (or directly σ_{s1}) are usually the pieces that comprise scattering data in input libraries for reactor calculations, while higher order Legendre moments need to be provided for specialized problems where more anisotropic scattering is expected.

If we define K_{N_s} by truncating the expansion (3.20) of its kernel at $n = N_s$, it follows from (3.21) and orthogonality of spherical harmonics that

$$K_{N_s}\psi = K\Pi_{P_N}\psi \tag{3.22}$$

provided that $N_s \leq N$. In other words, if physics of the scattering process allow it to be represented by an N_s-term expansion (3.21) where the *degree of scattering anisotropy* $N_s \leq N$, the P_N approximation will not introduce any additional error to the scattering source.

3.3.3 Rotational invariance of P_N equations

To expose rotational invariance of the P_N equations in the sense of Def. 2 (pg. 28), we will need some generally known facts about spherical harmonics that will also be of use later in Chap. 5.

Orthogonal decomposition of $L^2(\mathcal{S}_2)$

The $2n+1$ mutually orthogonal spherical harmonics of given degree n are the eigenfunctions of the Laplace operator on $\mathcal{S}_2$ corresponding to $\lambda_n = -n(n+1)$:

$$\nabla^2_{\mathcal{S}_2} Y_n^m(\boldsymbol{\Omega}) = -n(n+1)Y_n^m(\boldsymbol{\Omega}) \quad \forall -n \leq m \leq n$$

and generate the eigenspace

$$\Lambda_n = \mathrm{Span}\big\{Y_n^m; -n \leq m \leq n\big\}. \tag{3.23}$$

For $n = 0, 1, \ldots$, these finite-dimensional eigenspaces are closed, mutually orthogonal subspaces of $L^2(\mathcal{S}_2)$ and $\cup_{n=0}^{\infty}\Lambda_n$ is dense in $L^2(\mathcal{S}_2)$ ([49]), so that we have

$$L^2(\mathcal{S}_2) = \bigoplus_{n=0}^{\infty} \Lambda_n. \tag{3.24}$$

Restricting to a finite direct sum, we can hence write the P_N projection (3.13) as

$$\Pi_{P_N}\psi = \sum_{n=0}^{N} \Pi_{\Lambda_n}\psi, \tag{3.25}$$

where $\Pi_{\Lambda_n} : L^2(\mathcal{S}_2) \to \Lambda_n$, defined by

$$\big(\Pi_{\Lambda_n}\psi\big)(\mathbf{\Omega}) = \sum_{m=-n}^{n} \big(\psi, Y_n^m\big)_{L^2(\mathcal{S}_2)} Y_n^m(\mathbf{\Omega}),$$

is the orthogonal projection onto Λ_n.

Each Λ_n is a Hilbert space with the following *reproducing kernel property*:

Lemma 4. *For every $f \in \Lambda_n$, $n = 0, 1, \ldots$,*

$$f(\mathbf{\Omega}) = \frac{2n+1}{4\pi} \int_{\mathcal{S}_2} f(\mathbf{\Omega}')P_n(\mathbf{\Omega}\cdot\mathbf{\Omega}')\, d\mathbf{\Omega}'\,.$$

Proof. Follows from the addition theorem (A.5) and orthogonality of spherical harmonics. If $f \in \Lambda_n$,

$$\begin{aligned}\frac{2n+1}{4\pi} \int_{\mathcal{S}_2} f(\mathbf{\Omega}')P_n(\mathbf{\Omega}\cdot\mathbf{\Omega}')\,\mathrm{d}\mathbf{\Omega} &= \int_{\mathcal{S}_2} f(\mathbf{\Omega}') \sum_{j=-n}^{n} Y_n^j(\mathbf{\Omega})Y_n^j(\mathbf{\Omega}')\,\mathrm{d}\mathbf{\Omega} \\ &= \sum_{j=-n}^{n} \left(\int_{\mathcal{S}_2} f(\mathbf{\Omega}')Y_n^j(\mathbf{\Omega}')\,\mathrm{d}\mathbf{\Omega} \right) Y_n^j(\mathbf{\Omega}) = \Pi_{\Lambda_n} f(\mathbf{\Omega}) = f(\mathbf{\Omega}).\end{aligned}$$

□

Using the decomposition (3.25), we immediately obtain the following

Corollary 3. *For every $f \in L^2_K(\mathcal{S}_2)$,*

$$f(\mathbf{\Omega}) = \sum_{n=0}^{N} \frac{2n+1}{4\pi} \int_{\mathcal{S}_2} f(\mathbf{\Omega}')P_n(\mathbf{\Omega}\cdot\mathbf{\Omega}')\, d\mathbf{\Omega}'\,. \tag{3.26}$$

Now we are ready to show that the P_N equations are rotationally invariant in cases when the original NTE is. We note that even though the rotation operator $\mathcal{R}$ acts on both the spatial and angular variables, the P_N projection operator acts only on the latter. Therefore, we can still consider ψ as a function of only the angular variable in the proof of the following theorem. It should be also mentioned that even though rotational invariance of P_N equations is generally known, we could not find a formal proof of the fact in available literature, which served as a motivation for the theorem.

Theorem 4. *Let T be the transport operator defined in Sec. 2.2.5 such that the assumptions of Theorem 3 are satisfied. Then the corresponding P_N operator*

$$\Pi_{P_N} T \Pi_{P_N}$$

satisfies

$$\mathcal{R}\Pi_{P_N} T \Pi_{P_N} = \Pi_{S_N} T \Pi_{S_N} \mathcal{R} \quad \forall \mathcal{R} \in \mathrm{SO}(3).$$

Proof. Because of the commutativity of T and $\mathcal{R}$, it suffices to show that Π_{P_N} commutes with $\mathcal{R}$, that is

$$\mathcal{R}\Pi_{P_N}\psi = \Pi_{P_N}\mathcal{R}\psi \quad \forall \psi \in L^2(\mathcal{S}_2). \tag{3.27}$$

Noticing that $\Pi_{P_N}\psi \in L^2_K(\mathcal{S}_2)$ and using Corollary 3, we have

$$\begin{aligned}\mathcal{R}\Pi_{P_N}\psi &= \sum_{n=0}^{N} \frac{2n+1}{4\pi} \int_{\mathcal{S}_2} \psi(\mathbf{\Omega}') P_n(\mathbf{R}^T\mathbf{\Omega}\cdot\mathbf{\Omega}')\,\mathrm{d}\mathbf{\Omega}' \\ &= \sum_{n=0}^{N} \frac{2n+1}{4\pi} \int_{\mathcal{S}_2} \psi(\mathbf{\Omega}') P_n(\mathbf{\Omega}\cdot\mathbf{R}\mathbf{\Omega}')\,\mathrm{d}\mathbf{\Omega}'\,.\end{aligned}$$

Under substitution $\mathbf{\Omega}'' = \mathbf{R}\mathbf{\Omega}'$ with unit Jacobian determinant (since $\mathbf{R}$ is an orthogonal matrix), the expression becomes

$$\mathcal{R}\Pi_{P_N}\psi = \sum_{n=0}^{N} \frac{2n+1}{4\pi} \int_{\mathcal{S}_2} \psi(\mathbf{R}^T\mathbf{\Omega}'') P_n(\mathbf{\Omega}\cdot\mathbf{\Omega}'')\,\mathrm{d}\mathbf{\Omega}''\,.$$

After changing double primes back to single primes, this expression is equal to

$$\Pi_{P_N}\mathcal{R}\psi = \sum_{n=0}^{N} \frac{2n+1}{4\pi} \int_{\mathcal{S}_2} \psi(\mathbf{R}^T\mathbf{\Omega}') P_n(\mathbf{\Omega}\cdot\mathbf{\Omega}')\,\mathrm{d}\mathbf{\Omega}'\,.$$

□

3.3.4 Drawbacks of the P_N approximation

Using the results of the preceding subsection and well-known results from the theory of Hilbert spaces, we can see that the sum (3.25) (or (3.9)) converges in the $L^2(\mathcal{S}_2)$ norm to the true solution of eq. (2.6) as $N \to \infty$ ([49, Thm.

3.54]). However, the convergence may be very slow if the true solution to the NTE is not sufficiently regular in angle. In particular, pointwise convergence is hindered in the neighborhood of phase space points where the exact solution has jump discontinuity in $\mathbf{\Omega}$ (which may occur for example when a narrow beam of neutrons is freely streaming through a non-interacting medium, but, as we already mentioned at the beginning of Sec. 2.2.4, also in a more typical case of domains with multiple regions with different materials, bounded by piecewise polygonal boundary) and spurious oscillations are introduced to the approximate solution at these points. These oscillations spread over the whole angular domain and slow down the norm-wise convergence. This is a well-known property of Fourier series known as *Gibbs phenomenon*. Moreover, these oscillations do not vanish as more terms in the series are retained.

There are several ways of circumventing the Gibbs phenomenon. For example, when considering (3.9) as a means of deriving the P_N system, we may note that using a finite expansion obtained by truncating (3.9) at $n = N$ is not the only way of obtaining a closed system of equations – different closures are possible as we have already mentioned before. This fact has been utilized in [84] where the expansion has been adjusted to mitigate the oscillations by controlling angular gradients[9]. For other similar approaches in the context of general spectral methods, see e.g. [111].

As shown in [85], there is also another issue connected with time-dependent P_N approximation that must be kept in mind particularly when solving coupled problems. This issue is inherent in the structure of the P_N system and cannot be completely removed without losing some of its attractive properties. Namely, the authors proved that without sources and reactions, the linear hyperbolic character of eq. (3.10) (with an additional time derivative term as in (B.1)) together with rotational invariance allows negative solutions for positive, isotropic data in more than one dimension. To prevent negative solutions, one could either give-up linearity (e.g. by using a non-linear closure in a similar way as described above), rotational invariance (thus possibly introducing a different source of spurious oscillations plaguing the S_N method – the *ray-effects* discussed in Sec. 3.4.1.1) or

[9] Although (3.25) represents the best $L^2(\mathcal{S}_2)$ approximation of ψ by spherical polynomials up to given degree, absence of angular gradients in $L^2(\mathcal{S}_2)$ norm permits arbitrary oscillations.

hyperbolicity (thus changing the speed at which radiation propagates throughout the domain) – none of which is a generally satisfactory remedy. The authors also demonstrated that negative solutions can appear even in heterogeneous domains containing regions with reactions or sources.

3.3.5 Diffusion approximation

The set of monoenergetic steady state P_1 equations, obtained by assuming only linear angular variation of neutron flux:

$$\begin{aligned}\psi(\mathbf{r},\mathbf{\Omega}) &\approx \phi_0^0(\mathbf{r})Y_0^0(\mathbf{\Omega}) + \phi_1^{-1}(\mathbf{r})Y_1^{-1}(\mathbf{\Omega}) + \phi_1^0(\mathbf{r})Y_1^0(\mathbf{\Omega}) + \phi_1^1(\mathbf{r})Y_1^{-1}(\mathbf{\Omega}) \\ &= \sqrt{\tfrac{1}{4\pi}}\phi(\mathbf{r}) + \mathbf{\Omega}\cdot\mathrm{J}(\mathbf{r})\end{aligned}$$

can be under an additional assumption of vanishing anisotropic moments of sources (i.e. $q_k = 0,\ k = 1,2,\ldots$) and nonzero total cross-section (i.e., outside void regions) recast into a single elliptic equation[10]:

$$\begin{gathered}-\nabla\cdot D(\mathbf{r})\nabla\phi(\mathbf{r}) + \big[\sigma_t(\mathbf{r}) - \sigma_{s0}(\mathbf{r}) - \nu\sigma_f(\mathbf{r})\big]\phi(\mathbf{r}) = q_0(\mathbf{r}), \\ D(\mathbf{r}) \coloneqq \frac{1}{3\left[\sigma_t(\mathbf{r}) - \sigma_{s1}(\mathbf{r})\right]}\end{gathered} \tag{3.28}$$

with (cf. Sec. 3.3.2.4)

$$\sigma_{sn}(\mathbf{r}) = 2\pi\int_{-1}^{1}\sigma_s(\mathbf{r},\mu_0)P_n(\mu_0)\mathrm{d}\mu_0\,, \quad \mu_0 \equiv \mathbf{\Omega}\cdot\mathbf{\Omega}', \tag{3.29}$$

and appropriate form of the Marshak boundary conditions. Being mostly used for reactor criticality calculations, it is usually associated with the homogeneous boundary conditions of type (2.8) (including the vanishing ψ_{in} for $\beta = 0$), which in the Marshak approximation read

$$\mathbf{n}(\mathbf{r})\cdot D(\mathbf{r})\nabla\phi(\mathbf{r}) + \gamma(\mathbf{r})\phi(\mathbf{r}) = 0, \quad \gamma(\mathbf{r}) = \frac{1-\beta(\mathbf{r})}{2(1+\beta(\mathbf{r}))}, \quad \mathbf{r}\in\partial\mathcal{D}. \tag{3.30}$$

[10]We tacitly assume here that $D\nabla\phi$ is differentiable in $\mathcal{D}$. This unrealistic regularity assumption is relaxed in practice when we look for a weak solution of (3.28); we postpone the weak formulation of (3.28–3.30) to Sec. 3.5.3, but note that it could be also derived directly from the weak form of the P_1 equations (3.17),(3.19) with $q_1 \equiv 0$.

Note that neutron current is given in this approximation by:

$$\mathrm{J}(\mathbf{r}) = -D(\mathbf{r})\nabla\phi(\mathbf{r})$$

and eq. (3.28) can be also derived from physical conservation principles by employing the Fick's law of diffusion ([109]).

Equations (3.28–3.30) comprise the familiar neutron diffusion approximation. Thanks to its simplicity and also the efficiency of numerical solution techniques available for this approximation, it has always served as a "workhorse computational method of nuclear reactor physics" [109, p. 43]. The model is indeed sufficiently accurate for whole core calculations of contemporary reactors, assuming that the significant finer-scale neutron transport processes have been resolved by higher-fidelity NTE solvers applied in previous solution stages (as discussed in Sec. 3.2.1). The self-adjoint diffusion equation can then be solved using e.g. the finite element method in conjunction with both powerful and theoretically well-established conjugate gradient method with symmetric preconditioners like the algebraic multigrid ([16, 55]). Solution efficiency may be improved even further by using adaptive mesh refinement based on highly developed a posteriori error estimates for elliptic problems ([35, 54, 108]). Note that the self-adjoint property of the diffusion model can only be spoiled by the multigroup energy discretization, where energy transfers in neutron collisions result in non-symmetric coupling of the multigroup system – as we have seen before (Sec. 3.1), this can be prevented by moving the non-symmetric parts to the right-hand side and solving the resulting system iteratively.

Although methods based on diffusion approximation have been experimentally proven to be widely applicable for nuclear reactor analyses, there are situations where this approximation is just too coarse and, as some recent reports indicate [24, 40], these cases are likely to grow soon with the advent of new reactor and fuel designs. This approximation, of course, can also be hardly expected to produce acceptable results for more general problems with strong transport effects, where its basic assumptions are violated. One possibility then is to treat the diffusion solves not as a means of obtaining the final solution, but as *preconditioning* of an iteration involving a rigorous transport solution. Particularly popular became

such coupling between the diffusion calculation and discrete ordinates source iteration, which we return to in Sec. 3.4.3.

As another approach, we may try to generalize the procedure used to obtain the diffusion equation from the zeroth and first moment equations of the P_1 set to higher order P_N systems. Although this leads to an attractive system of weakly coupled diffusion-reaction equations in 1D, a complicated system of strongly coupled equations with mixed second-order partial derivatives results in more dimensions ([19]). To circumvent the problem, the *simplified P_N approximation* (SP_N) has been constructed by E. Gelbard [51, 52]) in the 1960's. This approximation will be the subject of Chapter 4.

3.4 The S_N method

Let us now turn our attention to the S_N approximation. The standard derivation uses the collocation approach in which a set of directions (*ordinates*) $\omega = \{\mathbf{\Omega}_m\}_{m=1}^M$ is chosen and the solution is approximated as:

$$\psi(\mathbf{r}, \mathbf{\Omega}) \approx \begin{cases} \psi(\mathbf{r}, \mathbf{\Omega}_m) & \text{if } \mathbf{\Omega} = \mathbf{\Omega}_m \text{ with } \mathbf{\Omega}_m \in \omega, \\ 0 & \text{if } \mathbf{\Omega} \notin \omega. \end{cases} \tag{3.31}$$

Equation (2.6) as well as the boundary conditions (2.7) or (2.8) are then evaluated at these $M = M(N)$ isolated directions. Notice that reflective (or albedo) boundary conditions place restrictions on the set of ordinates as it should optimally contain both directions of each reflected pair (otherwise an interpolation is needed). For the traditional direction sets, we have $M = N(N+2)$ if the given problem does not have any symmetries; the method of discrete ordinates using such a number of directions is traditionally referred to as the method of discrete ordinates of order N, shortly S_N.

In order to evaluate the integral term on the right hand side of the NTE, the set of directions is accompanied by a corresponding set of weights $\mathcal{W} = \{w_m\}_{m=1}^M$, together defining a quadrature of the sphere $\mathcal{S}_2$. The requirement of accurate evaluation of the integral term as well as accurate integration of angular flux over the sphere (to obtain the scalar flux) provides the main guideline for the choice of directions and weights. We will return to this matter later in Sec. 3.4.4; for now

it suffices to say that for three-dimensional problems without any symmetries, $M = |\{\mathbf{\Omega}_m, w_m\}| = \mathcal{O}(N^2)$ (with the value of M for typically used quadrature sets stated above).

To write the final result of the S_N approximation, let us first recall that for a sequence $s = \{c_k\}_N$, col s denotes the column vector with entries $c_1, c_2, \ldots, c_N$ and diag s the diagonal matrix defined by the elements of s. We define the vector functions representing S_N solution and sources, respectively, as

$$\Psi(\mathbf{r}) \coloneqq \operatorname{col}\{\psi_m(\mathbf{r})\}_M, \quad \mathrm{Q}_{S_N}(\mathbf{r}) \coloneqq \operatorname{col}\{q_m(\mathbf{r})\}_M, \tag{3.32}$$

[11]

the components of which are the fluxes and sources in ordinate directions

$$\psi_m(\mathbf{r}) \equiv \psi(\mathbf{r}, \mathbf{\Omega}_m), \quad q_m(\mathbf{r}) \equiv q(\mathbf{r}, \mathbf{\Omega}_m), \quad m = 1, \ldots, M. \tag{3.33}$$

The S_N approximation consists of the following set of M spatial PDEs:

$$\mathbf{A}^x_{S_N} \frac{\partial \Psi(\mathbf{r})}{\partial x} + \mathbf{A}^y_{S_N} \frac{\partial \Psi(\mathbf{r})}{\partial y} + \mathbf{A}^z_{S_N} \frac{\partial \Psi(\mathbf{r})}{\partial z} + \big[\sigma_t(\mathbf{r})\mathbf{I} - \mathbf{K}_{S_N}(\mathbf{r})\big]\Psi(\mathbf{r}) = \mathrm{Q}_{S_N}(\mathbf{r}), \tag{3.34}$$

where $\mathbf{r} \in \mathcal{D}$, $\mathbf{I}$ is the $M \times M$ identity matrix,

$$\mathbf{A}^x_{S_N} = \operatorname{diag}\{\Omega_{mx}\}_M, \ \mathbf{A}^y_{S_N} = \operatorname{diag}\{\Omega_{my}\}_M, \ \mathbf{A}^z_{S_N} = \operatorname{diag}\{\Omega_{mz}\}_M.$$

and

$$[\mathbf{K}_{S_N}(\mathbf{r})]_{m,n} = w_n \kappa(\mathbf{r}, \mathbf{\Omega}_m \cdot \mathbf{\Omega}_n), \quad \mathbf{\Omega}_n, \mathbf{\Omega}_m \in \omega, \ w_n \in \mathcal{W}, \ 1 \le m, n \le M. \tag{3.35}$$

After solution vector $\Psi(\mathbf{r})$ has been computed, the important physical quantities (scalar flux, neutron current) can be obtained directly from their definition (2.16) (2.17) using the quadrature formula, e.g.

$$\begin{aligned} \phi(\mathbf{r}) &= \int_{\mathcal{S}_2} \psi(\mathbf{r}, \mathbf{\Omega}) \, \mathrm{d}\mathbf{\Omega} \\ &= \sum_{m=1}^{M} w_m \psi(\mathbf{r}, \mathbf{\Omega}_m) = \sum_{m=1}^{M} w_m \psi_m(\mathbf{r}), \quad \mathbf{r} \in \overline{\mathcal{D}}, \ \mathbf{\Omega}_m \in \omega, \ w_m \in \mathcal{W}. \end{aligned} \tag{3.36}$$

[11] Using the same letter for both the S_N sources and the P_N source moments should not cause confusion, as the two will not show up at the same place.

3.4.1 Structure of the S_N approximation

3.4.1.1 Advection

Equation (3.34) represents a system of advection-reaction equations, each with constant advection field given by the matrices $\mathbf{A}^x_{S_N}, \mathbf{A}^y_{S_N}, \mathbf{A}^z_{S_N}$. We can see that unlike the P_N case, it has the form of a decoupled hyperbolic system, having M unique plane-wave solutions propagating in directions $\{\boldsymbol{\Omega}_m\}_M$. A plane-wave propagating in direction $\mathbf{R}^T\boldsymbol{\Omega}_m$, where $\mathbf{R}$ is the matrix representation of rotation $\mathcal{R} \in \mathrm{SO}(3)$, will be a solution to the S_N equations only if $\mathbf{R}^T\boldsymbol{\Omega}_m \in \omega$. We can therefore see that the fundamental property of rotational invariance (in case of rotationally invariant input data, see Sec. 2.2.7) is lost when approximating the continuous NTE by a finite S_N system.

In practice, this undesirable property manifests itself in the form of so-called *ray effects*. As a consequence of radiation propagating in a finite set of distinct directions, there will remain under-treated regions of the phase space, while other regions will receive more radiation in order to satisfy the global balance. This leads to spurious spatial oscillations of scalar flux, which become more pronounced as the spatial discretization is refined. This issue is somewhat ameliorated when strong scattering is present (as it increases the coupling of the equations, though it also makes them more difficult to solve numerically – see Sec. 3.4.3.1), but its true nature shows that the only systematic way of reducing ray effects in the framework of the S_N method is to increase the number of ordinates (especially when trying to reduce spatial discretization errors by using finer mesh). Being arguably the biggest issue of the S_N approximation, various more practical remedies have been proposed in literature (which, as expected, typically sacrifice some properties of the S_N approximation), see e.g. [59] and the references therein.

3.4.1.2 Boundary conditions

The decoupled hyperbolic character allows straightforward determination of inflow and outflow boundaries. Boundary conditions (2.7) are then easily taken into account as long as the ordinates set ω contains directions in which ψ_{in} is specified:

$$\psi_m(\mathbf{r}) = \psi_{\mathrm{in}}(\mathbf{r}, \boldsymbol{\Omega}_m) \quad \mathbf{r} \in \partial\mathcal{D},\ \boldsymbol{\Omega}_m \cdot \mathbf{n} < 0, \tag{3.37}$$

while (as already mentioned above), the conditions of type (2.8) require that for each $\mathbf{\Omega}_m \in \omega$,

$$\mathbf{\Omega}_{mR} \equiv \mathbf{\Omega}_m - 2\mathbf{n}(\mathbf{\Omega}_m \cdot \mathbf{n}) \in \omega.$$

This constraint is difficult to satisfy for generally oriented surfaces and some sort of interpolation of angular fluxes in directions adjacent to $\mathbf{\Omega}_{mR}$ is typically needed.

3.4.1.3 Collisions

Input data for the collision term usually include the isotropic fission cross-section and Legendre scattering moments up to some finite degree of scattering anisotropy N_s, as explained in Sec. 3.3.2.4. In order to incorporate this data to the S_N approximation, the expansion (3.20) truncated to length N_s is used in conjunction with the addition theorem (A.5) to obtain elements of the collision matrix:

$$[\mathbf{K}_{S_N}(\mathbf{r})]_{m,n} = w_n \sum_{p=0}^{N_s} \kappa_p(\mathbf{r}) \sum_{q=-p}^{p} Y_p^q(\mathbf{\Omega}_m) Y_p^q(\mathbf{\Omega}_n), \quad 1 \leq m, n \leq M, \quad \mathbf{r} \in \mathcal{D}. \tag{3.38}$$

3.4.1.4 Coupling of unknowns

In the P_N system, at most 7 unknowns are coupled by the advection term, while the collision term does not produce any coupling as a consequence of Lemma 3. On the contrary, the term $\mathbf{K}_{S_N}\Psi$ induces full unknown coupling as can be seen from (3.35) (while S_N advection matrices are diagonal). In order to recover sparsity (and also facilitate the use of efficient constant-advection solvers based on explicit marching in the advection direction), the so-called *source iteration* (SI) can be utilized, in which the system (3.34) is fully decoupled by moving $\mathbf{K}_{S_N}$ to the right hand side of S_N equations. Each equation is solved separately using any method suitable for an advection-reaction PDE with constant advection vector, using ψ_m from previous iteration to evaluate $\mathbf{K}_{S_N}\Psi$. Classical iteration methods like Jacobi and Gauss-Seidel are typically used to update Ψ during the iteration

process; e.g. the Jacobi scheme is given by the iteration

$$\mathbf{A}^x_{S_N}\frac{\partial\Psi_{(i+1)}}{\partial x}+\mathbf{A}^y_{S_N}\frac{\partial\Psi_{(i+1)}}{\partial y}+\mathbf{A}^z_{S_N}\frac{\partial\Psi_{(i+1)}}{\partial z}+\sigma_t\mathbf{I}\Psi_{(i+1)}=\mathbf{K}_{S_N}\Psi_{(i)}+\mathrm{Q}_{S_N},\ i=0,1,\ldots \tag{3.39}$$

for specified initial approximation $\Psi_{(0)}$.

In order to study convergence properties of this iteration in Sec. 3.4.3 by different means than the standard Fourier analysis (as presented e.g. in [3, Chap. III]), we will first represent the S_N method as a restriction of the original continuous NTE in an analogous way as in the case of P_N approximation.

3.4.2 Operator form of the S_N approximation

The aim of this subsection is to express the S_N system (3.34) in terms of the operators L and K from Chap. 2, function $\psi \in V$, and linear functional $q \in V'$, where V is a Hilbert space in which a unique solution of the NTE is ensured (cf. Sec. 2.2.5). To avoid technicalities involving dual spaces, we will again as in Sec. 2.2.5.3 consider $V = H^2_0(X)$ and $(L-K) : V \to L^2(X)$ and identify $q \in L^2(X) \subset V'$ with its Riesz representant in $L^2(X)$. As noted in Sec. 3.2.2, we further restrict our attention to the case of isotropic scattering, in which

$$\sigma_s(\cdot, \mathbf{\Omega}\cdot\mathbf{\Omega}') = \frac{\sigma_s}{4\pi}$$

so that

$$K\psi \equiv K_0\psi = \frac{\sigma_s+\nu\sigma_f}{4\pi}\int_{\mathcal{S}_2}\psi(\cdot,\mathbf{\Omega}')\,\mathrm{d}\mathbf{\Omega}' = \frac{\sigma_s+\nu\sigma_f}{4\pi}\phi.$$

Necessity of this restriction will be discussed in Sec. 3.4.2.1.

With this simplification in mind, let us consider an arbitrary S_N approximation of Problem 1, specified by the given set of ordinates $\omega = \{\mathbf{\Omega}_m\}_M$ and a corresponding set of quadrature weights $\mathcal{W} = \{w_m\}_M$. For each $\mathbf{\Omega}_m \in \omega$, we define a patch $\Delta\mathbf{\Omega}_m$ on the sphere with (spherical) area w_m (see Fig. 2.3 where the patch would correspond to the shaded area and $\Delta\mathbf{\Omega}_m = |\mathrm{d}\mathbf{\Omega}_m|$). Let ω define a complete covering of $\mathcal{S}_2$, so that

$$\sum_{m=1}^{M} w_m = \mu(\mathcal{S}_2) = 4\pi \quad \text{and} \quad w_m > 0,\ 1 \le m \le M. \tag{3.40}$$

This requirement is satisfied by most discrete ordinates quadrature sets in use. Taking into account the definition of the S_N unknowns by (3.32), let us define a mapping that transforms a function $u \in V$ to a spatial vector function F as:

$$\mathrm{F} = \widehat{\mathcal{I}}_{S_N} u = \mathrm{col}\,\{u(\cdot, \mathbf{\Omega}_m)\}_M\,.^{12} \tag{3.41}$$

Corresponding to it is the mapping which returns a function in V from a vector function $\mathrm{F} = \mathrm{col}\,\{f_m\}_M$:

$$\mathcal{I}_{S_N} : \mathrm{F} \mapsto u \in V_{S_N}, \quad u(\mathbf{r}, \mathbf{\Omega}) = \sum_{m=1}^{M} f_m(\mathbf{r}) \imath_m(\mathbf{\Omega}), \quad u|_{\partial X^-} = 0 \tag{3.42}$$

where $\imath_m$ is the indicator function of the patch:

$$\imath_m(\mathbf{\Omega}) = \begin{cases} 1 & \text{if } \mathbf{\Omega} \in \Delta\mathbf{\Omega}_m, \\ 0 & \text{otherwise.} \end{cases} \tag{3.43}$$

Here we introduced $V_{S_N} \subset V \subset L^2(X)$ as a subspace of a space of functions that are (as functions of $\mathbf{\Omega}$) piecewise constant on $\mathcal{S}_2$ and satisfy the homogeneous inflow boundary condition[13]. With these two mappings, we can now rewrite the S_N system (3.34) in terms of the transport operators from eq. (2.35):

$$\widehat{\mathcal{I}}_{S_N}(L - K_0)\mathcal{I}_{S_N}\Psi = \mathrm{Q}_{S_N}, \tag{3.44}$$

or, incorporating the definition of the S_N variables (3.32) by operator $\widehat{\mathcal{I}}_{S_N}$, as

$$\mathcal{I}_{S_N}\widehat{\mathcal{I}}_{S_N}(L - K_0)\mathcal{I}_{S_N}\widehat{\mathcal{I}}_{S_N}\psi = \mathcal{I}_{S_N}\widehat{\mathcal{I}}_{S_N} q \tag{3.45}$$

with $\psi \in V$, $q \in L^2(X)$. Finally, we can see from (3.41) and (3.42) that when $\widehat{\mathcal{I}}_{S_N}$ is restricted to V_{S_N}, $\mathcal{I}_{S_N}\widehat{\mathcal{I}}_{S_N} = I_{S_N}$ (identity on V_{S_N}) and thus the linear operator

$$\Pi_{S_N} \coloneqq \mathcal{I}_{S_N}\widehat{\mathcal{I}}_{S_N}, \tag{3.46}$$

is a projection $L^2(X) \to V_{S_N}$. The S_N system (3.34) can therefore be written as

$$\Pi_{S_N}(L - K_0)\Pi_{S_N}\psi = \Pi_{S_N} q, \tag{3.47}$$

that is, as a restriction of the original continuous NTE onto V_{S_N}.

[12] We implicitly include in this mapping the orthogonal projection onto a dense subspace of V comprising functions that are (as functions of $\mathbf{\Omega}$) continuous at $\mathbf{\Omega}_m \in \omega$ (this is a necessary technical step circumventing the problem of pointwise evaluation of functions $f(\mathbf{r}, \cdot) \in L^p(\mathcal{S}_2)$).

[13] Note that an analogous mapping could be defined for non-homogeneous boundary conditions and used in a lifting argument as in Sec. 2.2.5 to convert a non-homogeneous boundary value problem to the one analyzed here.

3.4.2.1 Remarks

It is worth noticing that even though the systems (3.44) and (3.34) (with (3.32), (3.33) and $K \equiv K_0$) are equivalent (in the sense that the solution vector Ψ of one system satisfies also the other provided the same source term q has been used), the interpretation of the solution of (3.44) is different from the point-wise approximation (3.31). The interpretation as a piecewise constant (w.r.t. to $\mathbf{\Omega}$) function over the ordinate patches is more natural, however, as it for example directly leads to the scalar flux definition (3.36).

It also deserves mentioning that Π_{S_N} is not an orthogonal projection in $L^2(X)$ as it is not symmetric on whole $\mathrm{Dom}\,(\Pi_{S_N})$. Therefore, the standard S_N approximation does not necessarily produce the best possible approximation of ψ by piecewise constant functions on $\mathcal{S}_2$ spanned by $\{\imath_m\}_M$. In order to obtain an orthogonal projection, the mapping $\widehat{\mathcal{I}}_{S_N}$ would have to be changed to

$$\widehat{\widehat{\mathcal{I}}}_{S_N} f = \mathrm{col}\left\{ \frac{1}{\Delta\mathbf{\Omega}_m} \int_{\Delta\mathbf{\Omega}_m} f(\cdot,\mathbf{\Omega})\imath_m(\mathbf{\Omega})\,\mathrm{d}\mathbf{\Omega} \right\}_M .$$

System (3.44) with this operator instead of $\widehat{\mathcal{I}}_{S_N}$ could be put into same form as the S_N system (3.34), but it would not be equivalent. To see this, define

$$\widehat{\Pi}_{S_N} = \mathcal{I}_{S_N}\widehat{\widehat{\mathcal{I}}}_{S_N}, \quad \psi_n = \left[\widehat{\widehat{\mathcal{I}}}_{S_N}\psi\right]_n,$$

split $L = A + \Sigma_t$ and notice that

$$\begin{aligned}
\widehat{\Pi}_{S_N} A \widehat{\Pi}_{S_N}\psi(\mathbf{r},\mathbf{\Omega}) &= \sum_{n=1}^{M}\left[\frac{1}{\Delta\mathbf{\Omega}_n}\int_{\Delta\mathbf{\Omega}_n} \mathbf{\Omega}\cdot\nabla \sum_{m=1}^{M}\psi_m(\mathbf{r})\imath_m(\mathbf{\Omega})\imath_n(\mathbf{\Omega})\,\mathrm{d}\mathbf{\Omega}\right]\imath_n(\mathbf{\Omega}) \\
&= \sum_{n=1}^{M}\left[\frac{1}{\Delta\mathbf{\Omega}_n}\int_{\Delta\mathbf{\Omega}_n} \mathbf{\Omega}\cdot\nabla\psi_n(\mathbf{r})\imath_n(\mathbf{\Omega})\,\mathrm{d}\mathbf{\Omega}\right]\imath_n(\mathbf{\Omega}) \\
&= \sum_{n=1}^{M}\left[\widehat{\mathcal{I}}_{S_N}\mathbf{\Omega}\right]_n\cdot\nabla\psi_n(\mathbf{r})\imath_n(\mathbf{\Omega});
\end{aligned}$$

that is, the directional derivative has to be considered with respect to the *average ordinate vector* $\mathbf{\Omega}_n = \left[\widehat{\widehat{\mathcal{I}}}_{S_N}\mathbf{\Omega}\right]_n$. Nevertheless, such a construction corresponds to another widely used numerical approximation translated to angular domain.

Writing the weak form of (3.47) with Π_{S_N} replaced by $\widehat{\Pi}_{S_N}$ and using symmetry of this modified projection operator, we get

$$\Big((L-K)\widehat{\Pi}_{S_N}\psi, \widehat{\Pi}_{S_N}\varphi\Big)_{L^2(X)} = \Big(q, \widehat{\Pi}_{S_N}\varphi\Big)_{L^2(X)} \quad \forall\varphi \in L^2(X). \tag{3.48}$$

This is a system obtained by applying in angular domain the discontinuous Galerkin (DG) method with piecewise constant shape functions generating $\operatorname{Range}\widehat{\Pi}_{S_N}$ $(= \operatorname{Range}\Pi_{S_N} = V_{S_N})$.

Problem analogous to that described in the previous paragraph also lies behind the restriction to isotropic scattering, for otherwise (neglecting the isotropic fission part of K)

$$\begin{aligned}\big(K\Pi_{S_N}\big)(\mathbf{r},\boldsymbol{\Omega}) &= \int_{\mathcal{S}_2} \sigma_s(\mathbf{r}, \boldsymbol{\Omega}\cdot\boldsymbol{\Omega}') \sum_{m=1}^{M} \psi_m(\mathbf{r})\imath_m(\boldsymbol{\Omega}')\,\mathrm{d}\boldsymbol{\Omega}' \\ &= \sum_{m=1}^{M} \psi_m(\mathbf{r}) \int_{\mathcal{S}_2} \sigma_s(\mathbf{r}, \boldsymbol{\Omega}\cdot\boldsymbol{\Omega}')\imath_m(\boldsymbol{\Omega}')\,\mathrm{d}\boldsymbol{\Omega}'\end{aligned}$$

and $\sigma_s(\mathbf{r}, \boldsymbol{\Omega}\cdot\boldsymbol{\Omega}') = \sigma_s(\mathbf{r})$ is a sufficient condition for the last integral to be equal to $w_m\sigma_s(\mathbf{r}, \boldsymbol{\Omega}\cdot\boldsymbol{\Omega}_m)$ (as needed for the equivalence with (3.35)). Note that this condition would also be satisfied if $\imath_m$ was an appropriate Dirac delta distribution in $\boldsymbol{\Omega}$, in which case, however, it would not belong to $L^2(\mathcal{S}_2)$.

3.4.3 Convergence of source iteration

Having established the connection between the fully continuous NTE on the Hilbert space $V = H_0^2(X)$ and its S_N approximation on $V_{S_N} \subset V$ in Sec. 3.4.2, we can now use properties of operators L and K_0 to investigate convergence of source iteration in the discrete ordinates approximation. To this end, let us first write (3.39) as an iteration on V_{S_N}:

$$\Pi_{S_N} L \Pi_{S_N} \psi_{(i+1)} = \Pi_{S_N} K_0 \Pi_{S_N} \psi_{(i)} + \Pi_{S_N} q, \quad i = 0, 1, \ldots \tag{3.49}$$

(assuming given $\psi_{(0)} \in V$). Again, this is just a restriction to V_{S_N} of the following iteration on $L^2(X)$:

$$L\psi_{(i+1)} = K_0\psi_{(i)} + q \quad i = 0, 1, \ldots. \tag{3.50}$$

Egger and Schlottbom [43] used Banach fixed point theory and the iteration (3.50) to show that the mapping

$$\mathcal{T}_q : u \mapsto L^{-1}Ku + L^{-1}q, \quad q \in L^p(X)$$

(with a general collision operator not restricted to isotropic scattering) is contractive for all $1 \leq p \leq \infty$ provided that conditions equivalent to subcriticality conditions (Def. 1) hold. With this result, the authors showed well-posedness of Problem 1. They also exhibited the contraction factor, which in our notation with c and d given by (2.23) and (2.25), respectively, can be written as:

$$\rho_p = 1 - e^{-C_p}, \quad C_p = \frac{1}{p}\|c\sigma_t\ell\|_{L^\infty(X)} + \frac{p-1}{p}\|d\sigma_t\ell\|_{L^\infty(X)}, \quad 1 \leq p \leq \infty \quad (3.51)$$

where $\ell = \ell(\mathbf{r}, \mathbf{\Omega})$ is the length of the characteristic line segment passing through $\mathbf{r}$ in the direction $\mathbf{\Omega}$. With the current assumptions on energy independence and isotropic scattering,

$$c(\mathbf{r}) = d(\mathbf{r}) = \frac{\sigma_s(\mathbf{r}) + \nu\sigma_f(\mathbf{r})}{\sigma_t(\mathbf{r})}, \quad \mathbf{r} \in \mathcal{D}.$$

We also assume ℓ to be bounded a.e. in X and such that

$$\forall \mathbf{\Omega} \in \mathcal{S}_2 \; \exists s_0 : \; 0 < s_0 < \ell(\mathbf{r}, \mathbf{\Omega}) \;\text{ and }\; \mathbf{r}_0 = \mathbf{r} - s_0\mathbf{\Omega} \in \partial X^-$$

(see Fig. 2.5 in Sec. 2.2.1).

As the contraction property

$$\exists \rho \in [0, 1) : \|\mathcal{T}_q\, u_1 - \mathcal{T}_q\, u_2\|_V \leq \rho\|u_1 - u_2\|_V \quad \forall u_1, u_2 \in V$$

holds also for any u_1, u_2 in a subspace of V, we immediately get from the contraction principle (e.g., [39, Thm. 2.3.1]) the convergence result for the S_N source iteration (3.49) and hence also for (3.39).

Theorem 5. *Let $V = H_0^2(X)$ and $V_{S_N} \subset V$ the S_N approximation subspace and let the subcriticality conditions in $L^2(X)$ hold. Then for any $\Psi_{(0)}$ (corresponding to $\psi_{(0)} \in V_{S_N}$ via the mapping $\widehat{\mathcal{I}}_{S_N}$), the sequence of iterates $\{\Psi_{(i)}\}_{i=1}^\infty$ of iteration* (3.39) *converges to the unique solution Ψ^* of equation* (3.44) *(with Q_{S_N} corresponding to a $q \in L^2(X)$). Moreover,*

$$\|\Psi_{(i)} - \Psi^*\| \leq \frac{\rho_2^i}{1-\rho_2}\|\Psi_{(1)} - \Psi_{(0)}\|, \quad \|\Psi_{(i)} - \Psi^*\| \leq \frac{\rho_2}{1-\rho_2}\|\Psi_{(i)} - \Psi_{(i-1)}\|,$$

where

$$\rho_2 = 1 - e^{-\|c\sigma_t \ell\|_{L^\infty(X)}}. \tag{3.52}$$

Remark 10 (Multigroup case). This approach may be extended to the energy-dependent case once we recognize that the multigroup system (3.3) is again a restriction of the continuous NTE to a subspace of piecewise constant functions with respect to energy, using $\Pi_G \coloneqq \mathcal{I}_G \widehat{\mathcal{I}}_G$ where

$$\widehat{\mathcal{I}}_G \psi = \left\{ \frac{1}{\Delta E^g} \int_g \psi(\cdot,\cdot,E)\,\mathrm{d}E \right\}_G, \quad \mathcal{I}_G\{\psi^g\}_G = \sum_{g=1}^{G} \psi^g(\cdot,\cdot)\imath_g(E)$$

with obviously defined group indicator function $\imath_g$. The above results then apply to the Jacobi version of iteration (3.6), while the Gauss-Seidel form could be analyzed by splitting the continuous transport operator as

$$T = L + K_\downarrow + K_\uparrow,$$

with

$$K_\downarrow \psi(\mathbf{r},E) = \int_{E \le E'} \int_{\mathcal{S}_2} \kappa(\mathbf{r}, \boldsymbol{\Omega}\cdot\boldsymbol{\Omega}', E \leftarrow E')\psi(\mathbf{r},\boldsymbol{\Omega}',E')\,\mathrm{d}\boldsymbol{\Omega}'\,\mathrm{d}E',$$
$$K_\uparrow \psi(\mathbf{r},E) = \int_{E > E'} \int_{\mathcal{S}_2} \kappa(\mathbf{r}, \boldsymbol{\Omega}\cdot\boldsymbol{\Omega}', E \leftarrow E')\psi(\mathbf{r},\boldsymbol{\Omega}',E')\,\mathrm{d}\boldsymbol{\Omega}'\,\mathrm{d}E'.$$

3.4.3.1 Source iteration under diffusive conditions

We can see from (3.52) that ρ_2 gets closer to 1 and the convergence rate deteriorates as scattering becomes the dominating collision event and size of the domain gets larger. This means that the *escape probability* of neutrons gets lower and neutrons diffuse through the domain for a long time until they get absorbed (note that if we assume non-fissioning domain with $\sigma_f = 0$ and neglect inelastic scattering, $c \to 1$ with fixed σ_t implies $\sigma_a \to 0$, cf. (2.24) and (2.14)). Such conditions are referred to as *diffusive conditions* and the classical asymptotic analysis (see e.g. [42] or the overview [2] and references therein) tells us that the solution of the diffusion approximation tends to the solution of the NTE under these conditions.

More precisely, "diffusivity" of the medium is characterized by a small parameter $\varepsilon \ll 1$ such that $\varepsilon \to 0$ reflects the properties described above. The

parameter is also defined in such a way that when the terms of the NTE are scaled by ε and $\varepsilon \to 0$, the diffusion equation (3.28) is obtained up to terms of order $\mathcal{O}(\varepsilon^3)$. Hence, in cases when the source iteration converges slowly, it makes sense to *precondition* it by the solution of the diffusion problem, which is known as the *diffusion synthetic acceleration of SI*. For more details, we refer to [9, Chap. 1] or [3, Sec. III] (where also convergence characteristics of the discrete ordinates source iteration similar to Thm. 5 were obtained for homogeneous infinite medium with isotropic scattering by means of Fourier analysis).

3.4.4 Selection of ordinates and weights

In this subsection, we return to the question of the selection of the angular quadrature set $\{\mathbf{\Omega}_m, w_m\}_M$. The first constraint that we encountered was connected with the reflective boundary conditions. If these conditions are imposed for the problem at hand, for each direction $\mathbf{\Omega}_m \in \omega$ we should require

$$\mathbf{\Omega}_{mR} \equiv \mathbf{\Omega}_m - 2\mathbf{n}(\mathbf{\Omega}_m \cdot \mathbf{n}) \in \omega, \tag{3.53}$$

where $\mathbf{n}$ is the unit outward normal to the reflective boundary. As already mentioned in Sec. 3.4.1.2, this condition needs to be imposed approximately for generally oriented surfaces. The set ω in most widely adopted ordinate sets is constructed so as to preserve the symmetry of the eight octants of $\mathcal{S}_2$ with respect to $\pi/2$ rotations (such sets are usually called *level-symmetric*). Constraint (3.53) is then satisfied for boundaries parallel to Cartesian coordinate planes. Moreover, ordinates and the corresponding weights need to be specified only in the principal octant of the sphere; the same direction cosines with only an appropriately changed sign define corresponding points in the remaining octants (with the same quadrature weights).

The choice of quadrature set is further guided by the requirement of exact integration of angular integrals appearing in the S_N model. This can be represented by the so-called *moment conditions* ([61]):

$$\sum_{m=1}^{M} w_m(\Omega_{mx})^n = \sum_{m=1}^{M} w_m(\Omega_{my})^n = \sum_{m=1}^{M} w_m(\Omega_{mz})^n = \begin{cases} 0 & n \text{ odd},, \\ \frac{4\pi}{n+1} & n \text{ even}. \end{cases} \tag{3.54}$$

Note that the odd moment conditions are automatically satisfied for level-symmetric sets. For $n = 0$, we recover the condition (3.40). This is needed for correct integration of isotropic terms (like the scalar flux). Satisfaction of the conditions for $n = 1$ is required for correct determination of neutron current from S_N fluxes (cf. eq. (2.17)). Satisfaction of higher order moments has more subtle physical significance, e.g. (3.54) for $n = 2$ is needed so that the S_N solution obeys the diffusion limit [72]; it is also important when higher anisotropy degrees are present in the scattering integral, as we will explain next.

Consider the action of $\mathbf{K}_{S_N}$ on the S_N representation of $Y_r^s(\mathbf{\Omega})$ for some admissible index pair r, s:

$$\begin{aligned}\left[\mathbf{K}_{S_N}\widehat{\mathcal{I}}_{S_N}Y_r^s\right]_m &= \sum_{n=1}^{M} w_n \sum_{p=0}^{N_s} \kappa_p \sum_{q=-p}^{p} Y_p^q(\mathbf{\Omega}_m)Y_p^q(\mathbf{\Omega}_n)Y_r^s(\mathbf{\Omega}_n) \\ &= \sum_{p=0}^{N_s} \kappa_p \sum_{q=-p}^{p} Y_p^q(\mathbf{\Omega}_m) \sum_{n=1}^{M} w_n Y_p^q(\mathbf{\Omega}_n)Y_r^s(\mathbf{\Omega}_n).\end{aligned}$$

If the quadrature set is constructed so that products of spherical harmonics up to degree N_s are integrated exactly, the orthogonality property (A.4) is recovered by the last sum and we obtain

$$\mathbf{K}_{S_N}\widehat{\mathcal{I}}_{S_N}Y_r^s = \{\kappa_r Y_r^s(\mathbf{\Omega}_m)\}_{m=1}^{M} = \widehat{\mathcal{I}}_{S_N}KY_r^s$$

where the last equality follows from (3.21) and definition of $\widehat{\mathcal{I}}_{S_N}$.

The requirement of exact integration of spherical harmonics up to degree N_s is equivalent to exact integration of polynomials in $\mathbf{\Omega}_x$, $\mathbf{\Omega}_y$, $\mathbf{\Omega}_z$ of the same degree (as both sets generate the same polynomial space on $\mathcal{S}_2$, cf. Tab. 5.1 and Sec. 5.2 for more details), i.e. conditions (3.54). That is, if conditions (3.54) are satisfied through $m = 2N_s$, the product of spherical harmonics up to degree N_s and hence their orthogonality relation will be integrated exactly.

To obtain an optimal quadrature rule capable of integrating exactly polynomials of highest possible degree, we use the well-known property of Gauss quadrature: if the N abscissas (nodes) of an N-point Gauss quadrature rule for integrating functions over given interval with a particular weighting function are chosen as the roots of the polynomial of degree N from a set of polynomials

orthogonal over the same interval with the same weighting function, then the quadrature rule will be exact for all polynomials up to degree $2N-1$ (e.g., [96, Chap. 4]). Realizing that with $\mu = \cos\vartheta$ and $y = \cos\varphi$,

$$\int_{\mathcal{S}_2} \mathrm{d}\mathbf{\Omega} = \int_0^{\pi} \sin\vartheta \mathrm{d}\vartheta \int_0^{2\pi} \mathrm{d}\varphi = 2\int_{-1}^{1} \mathrm{d}\mu \int_{-1}^{1} \frac{1}{\sqrt{1-y^2}}\,\mathrm{d}y\,.$$

we come to the conclusion that the quadrature rule optimal with respect to the above conditions is the tensor product rule composed of the N_μ-point Gauss-Legendre rule in the polar direction (as Legendre polynomials are orthogonal over $[-1,1]$ with the unit weight function) and the N_φ-point Gauss-Chebyshev rule in the azimuthal direction (as Chebyshev polynomials are orthogonal over $[-1,1]$ with the weight $(1-y^2)^{-1/2}$) with $N_\mu = N_\varphi = N_s + 1$ points (so that it is exact for polynomials of degree $2N_s+1$). As weight functions for both rules are positive over $[-1,1]$, the weights are also positive for arbitrary N_μ, N_φ and consequently both 1D quadrature rules comprising the final product rule are convergent ([96, Chap. 4]).

3.4.4.1 Legendre-Chebyshev quadrature in the first octant

Because of the symmetry constraints, we actually have for the S_N quadrature $N_\mu = N$ and $N_y = N+2$, resulting in ordinates set ω with $M = N(N+2)$ directions over the unit sphere. The directions are arranged in the principal octant on $N/2$ levels of constant polar angle ϑ_l (and thus constant $\mu_l = \Omega_{lz} = \cos\vartheta_l$) with l directions at the l-th level.

Polar components

For $l = 1, 2, \ldots N/2$, μ_l are the nodes of the Gauss-Legendre rule, i.e. the $N/2$ positive roots of Legendre polynomial $P_N(\mu)$. The weights are given as

$$w_l^\mu = \frac{1}{(1-\mu_l^2)\left(\frac{dP_N(\mu)}{d\mu}\Big|_{\mu_l}\right)^2}.$$

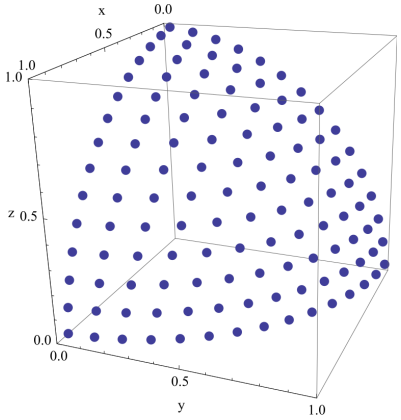


Figure 3.2: Legendre-Chebyshev ordinates in first octant, $N = 30$.

Azimuthal components

The azimuthal angle of the i-th direction on the l-th polar level is the arc-cosine of the i-th root of the Chebyshev polynomial of the first kind of degree l ([121, p.402])[14]. That is, for given $l = 1, 2, \ldots, N$,

$$\varphi_{l,i} = \frac{2l - 2i + 1}{2l} \times \frac{\pi}{2}, \quad i = 1, 2, \ldots, l.$$

All weights on given polar level l are equal to ([121, p.402])

$$w^{\varphi}_{l,i} = \frac{\pi}{l}, \quad i = 1, 2, \ldots, l.$$

Complete quadrature set

The Legendre-Chebyshev quadrature set

$$\{\mathbf{\Omega}_m, w_m\}_M \equiv \{\mathbf{\Omega}_{l,i}, w_{l,i} \mid l = 1, 2, \ldots N,\ i = 1, 2, \ldots l\}$$

is defined by

$$\mathbf{\Omega}_{l,i} = \begin{bmatrix} \sqrt{1-\mu_l^2}\cos\varphi_{l,i} \\ \sqrt{1-\mu_l^2}\sin\varphi_{l,i} \\ \mu_l \end{bmatrix}, \quad w_{l,i} = w^{\mu}_l w^{\varphi}_{l,i} \tag{3.55}$$

and is used in the S_N module for Hermes2D (Chap. 6). Mathematica script for generating the quadrature points and weights is available from `https://raw.githubusercontent.com/mhanus/hermes/SN-adaptive/hermes2d/examples/neutronics/ordinates.nb`.

[14] With the ordering of roots as in [121], more precisely the arc-cosine of the $l - i + 1$-st root.

3.5 Approximation of spatial dependence

3.5.1 S_N and P_N methods

As we have seen in previous sections, both the S_N and the P_N approximation leads to a system of linear hyperbolic PDE's in spatial variables. The final approximation step typically consists of laying out a mesh over the spatial domain and using finite difference (FD), finite volume (FV) or finite element (FE) methods to discretize the PDE's. In view of the Galerkin formulation of the P_N and S_N systems, it might be tempting to formulate the final restriction to a finite dimensional subspace of $H^2(X)$ in a consistent way, using as the projection target a subspace of $\operatorname{Range}\Pi_{S_N}$ or $\operatorname{Range}\Pi_{P_N}$ spanned by finite number of basis functions defined on $\mathcal{D}$. This can be in general done by using the finite element method, but it turns out that in current case, it is practical only if the P_N or S_N methods were applied on the second-order forms of the NTE.

We will explain the reason for the simpler case of the S_N approximation. We will also assume that the S_N equations were decoupled by the source iteration technique and consider a single step of the process (3.39) with all terms on the right grouped under the source term. Suppressing the iteration index, we may write the final system with vacuum boundary conditions as

$$\mathrm{L}_{S_N}\Psi = \mathrm{Q}_{S_N}, \quad \text{where} \quad \mathrm{L}_{S_N} \coloneqq \widehat{\mathcal{I}}_{S_N} L \mathcal{I}_{S_N},\ \mathrm{Q}_{S_N} \coloneqq \widehat{\mathcal{I}}_{S_N} q \tag{3.56}$$

or in the expanded form as:

$$\boldsymbol{\Omega}_m \cdot \nabla\psi_m(\mathbf{r}) + \sigma_t(\mathbf{r})\psi_m(\mathbf{r}) = q_m(\mathbf{r}), \quad \mathbf{r} \in \mathcal{D}, \tag{3.57}$$

$$\psi_m(\mathbf{r}) = 0, \qquad \mathbf{r} \in \partial\mathcal{D}_m^- \tag{3.58}$$

for $m = 1, 2, \ldots, M$, where

$$\partial\mathcal{D}_m^{\pm} = \{\mathbf{r} \in \partial\mathcal{D} : \boldsymbol{\Omega}_m \cdot \mathbf{n}(\mathbf{r}) \gtrless 0\}.$$

Let $\mathrm{U} = [u_1, u_2, \ldots, u_M]^T$, $\mathrm{V} = [v_1, v_2, \ldots, v_M]^T$ and $\mathbb{L}^2(\mathcal{D}) = \left[L^2(\mathcal{D})\right]^M$ with the inner product

$$(\mathrm{U}, \mathrm{V})_{\mathbb{L}^2(\mathcal{D})} = \int_{\mathcal{D}} \mathrm{U} \cdot \mathrm{V}\, \mathrm{d}\mathbf{r} = \sum_{m=1}^{M} \int_{\mathcal{D}} u_m(\mathbf{r}) v_m(\mathbf{r})\, \mathrm{d}\mathbf{r} = \sum_{m=1}^{M} (u_m, v_m)_{L^2(\mathcal{D})}.$$

Further, let

$$\mathbb{V}(\mathcal{D}) = \prod_{m=1}^{M} \mathcal{V}_m(\mathcal{D}), \quad \mathcal{V}_m(\mathcal{D}) = \{v \in L^2(\mathcal{D}) : \mathbf{\Omega}_m \cdot \nabla v + \sigma_t v \in L^2(\mathcal{D})\}.$$

The problem of finding a weak solution of (3.56) can now be formulated as a problem of finding $\Psi \in \mathbb{V}(\mathcal{D})$ with $\psi_m|_{\partial\mathcal{D}_m^-} = 0$ such that

$$a(\Psi, \mathrm{V}) = f(\mathrm{V}) \quad \forall \mathrm{V} \in \mathbb{L}^2(\mathcal{D}), \tag{3.59}$$

where

$$a(\Psi, \mathrm{V}) = (\mathrm{L}_{S_N}\Psi, \mathrm{V})_{\mathbb{L}^2(\mathcal{D})}, \quad f(\mathrm{V}) = (\mathrm{Q}_{S_N}, \mathrm{V})_{\mathbb{L}^2(\mathcal{D})}.$$

Expanding this weak formulation and imposing the inflow boundary conditions in the weak sense by using Green's theorem, we rewrite (3.59) as

$$\begin{cases} \displaystyle\sum_{m=1}^{M} a_{mm}(\psi_m, v_m) = \sum_{m=1}^{M} (q_m, v_m)_{L^2(\mathcal{D})} \quad \forall \mathrm{V} \in \mathbb{V}(\mathcal{D}), \\ \Psi \in \mathbb{V}(\mathcal{D}) \end{cases} \tag{3.60}$$

with

$$a_{mm}(u, v) = \int_{\mathcal{D}} (-u\mathbf{\Omega}_m \cdot \nabla v + \sigma_t uv)\, \mathrm{d}\mathbf{r} + \int_{\partial\mathcal{D}_m^+} uv\, \mathbf{\Omega}_m \cdot \mathbf{n} \mathrm{d}S\,.$$

Provided that (3.53) hold, reflective or albedo boundary conditions (2.8) are weakly imposed by adding to a_{mm} the following surface integral:

$$\int_{\partial\mathcal{D}_m^-} \psi_{mR} v_m\, \mathbf{\Omega}_m \cdot \mathbf{n} \mathrm{d}S\,, \quad \partial\mathcal{D} \cap \partial\mathcal{D}_{\text{reflective}} \neq \emptyset. \tag{3.61}$$

For fixed m, arbitrary $v_m \in \mathcal{V}_m(\mathcal{D})$ and $\mathrm{V} = [0, \ldots, v_m, \ldots, 0]^T$, we obtain from (3.60) the weak form of the m-th advection-reaction equation

$$\begin{cases} a_{mm}(\psi_m, v_m) = (q_m, v_m)_{L^2(\mathcal{D})} \quad \forall v_m \in \mathcal{V}_m(\mathcal{D}), \\ \psi_m \in \mathcal{V}_m(\mathcal{D}). \end{cases} \tag{3.62}$$

Let us further consider this single advection-reaction problem and suppress the index m.

3.5.2 Finite element method

We proceed by defining a (general unstructured, quasiuniform) mesh $\mathcal{T}_h = \{\tau\}$ (τ being either a simplex or a hypercube) such that

$$\overline{\mathcal{D}} = \bigcup_{\tau \in \mathcal{T}_h} \overline{\tau}$$

where h denotes the maximum diameter of $\tau \in \mathcal{T}_h$. The (conforming) finite element method restricts (3.62) to the finite-dimensional subspace of $\mathcal{V}$

$$\mathcal{V}_{hp} = \operatorname{Range} \Pi_{hp}\, u, \tag{3.63}$$

where the projection operator may be expressed (like the S_N or P_N projections) as $\Pi_{hp} = \mathcal{I}_{hp}\widehat{\mathcal{I}}_{hp}$, where

$$\widehat{\mathcal{I}}_{hp} : u \mapsto \mathbf{u} \in \mathbb{R}^{N_{hp}}, \quad \mathcal{I}_{hp} : \mathbf{u} \mapsto u_{hp} = \sum_{i=1}^{N_{hp}} u_i s_i(\mathbf{r}).$$

Here for $i = 1, \ldots, N_{hp}$, $s_i \in C^0(\mathcal{D})$ are the globally continuous *shape functions* generating $\mathcal{V}_{hp}$ and the mapping $\widehat{\mathcal{I}}_{hp}$ is defined by the choice of finite elements type (see [78, Chap. 3], [90]). The classical choice of shape functions are the piecewise polynomial functions, giving rise to the approximation subspace

$$\mathcal{V}_{hp} = \{v_{hp} \in C^0(\mathcal{D}) : v_{hp}|_\tau \circ \mathfrak{r} \in \mathcal{P}_p(\hat{\tau}),\ \tau \in \mathcal{T}_h\} \tag{3.64}$$

where $\hat{\tau}$ is the reference unit hypercube or simplex, $\mathfrak{r} : \tau \to \hat{\tau}$ is the standard reference mapping and $\mathcal{P}_p(\hat{\tau})$ is the space of polynomials of degree up to p (tensor product polynomials in case of τ being a hypercube).

As a result of the restriction of (3.62) to $\mathcal{V}_{hp}$, we obtain for $u, v \in \mathcal{V}$

$$a(u_{hp}, v_{hp}) = a(\Pi_{hp}u, \Pi_{hp}v) = a(\mathcal{I}_{hp}\mathbf{u}, \mathcal{I}_{hp}\mathbf{v}) = (\mathbf{A}\mathbf{u})^T\mathbf{v} \tag{3.65}$$

where

$$[\mathbf{A}]_{ij} = (\mathbf{A}\mathbf{e}_i)^T\mathbf{e}_j = a(s_i, s_j)$$

and analogously

$$(q, \Pi_{hp}v) = \mathbf{b}^T\mathbf{v}, \quad [\mathbf{b}]_i = (q, s_i)_{L^2(\mathcal{D})};$$

that is, the algebraic system

$$\mathbf{A}\mathbf{u} = \mathbf{b}, \quad \mathbf{A} \in \mathbb{R}^{N_{hp} \times N_{hp}}, \ \mathbf{u}, \mathbf{b} \in \mathbb{R}^{N_{hp}}.$$

Returning back to the general case with $m = 1, \ldots, M$, the matrix $\mathbf{A}$ would form the mm block in the global S_N matrix (restriction of eq. (3.59) to $\prod_m \mathcal{V}_m^{hp}(\mathcal{D})$).

By this construction, $\mathcal{V}_{hp} \subset \mathcal{V}$ [15] and we may consider the above procedure as another restriction of the NTE following its restriction to V_{SN}. The finite dimensional subspace in which we are looking for the solution has dimension $N_{hp} \times M \times G$ (if we also consider the multigroup discretization of energy).

Equations of type (3.62) have been studied extensively in the past (see e.g. [56] and references therein) and it is well known that the bilinear form in (3.62) is not coercive on $\mathcal{V}_m(\mathcal{D})$. Namely, there exists $\alpha > 0$ such that for any $v \in \mathcal{V}_m(\mathcal{D})$,

$$a_{mm}(v, v) \geq \alpha \|v\|^2_{L^2(\mathcal{D})} + \frac{1}{2} \int_{\partial\mathcal{D}} |\mathbf{\Omega}_m \cdot \mathbf{n}| \, v^2 \, \mathrm{d}S \,, \tag{3.66}$$

while the norm on $\mathcal{V}_m(\mathcal{D})$ is given by

$$\|v\|^2_{\mathcal{V}_m(\mathcal{D})} = \|\mathbf{\Omega}_m \cdot \nabla v\|^2_{L^2(\mathcal{D})} + \|v\|^2_{L^2(\mathcal{D})} + \int_{\partial\mathcal{D}} |\mathbf{\Omega}_m \cdot \mathbf{n}| \, v^2 \, \mathrm{d}S \,.$$

As coercivity is preserved by restricting to a subspace, this property may be expected to be reflected by the discretization described above. Indeed, if the exact solution is not sufficiently smooth (and it is generally not in practice) unstable results with spurious oscillations arise because the directional derivative is uncontrolled by the right-hand side of (3.66).

3.5.2.1 Discontinuous Galerkin method

To circumvent this issue, one approach is to give up conformity and define an appropriate approximation space on which the restricted bilinear form can be shown to be coercive. This leads to the *discontinuous Galerkin method* of order

[15] provided that (3.63) holds, which is often violated by the inability to precisely capture the geometrical boundaries; for more details on this as well as other possible *variational crimes*, we refer to [107, Sec. 4.1.2]

p, shortly DG(p), where the shape functions are still polynomials of degree up to p on each element, but are not required to be globally continuous:

$$\mathcal{V}_{hp}^{dg} = \{v_{hp} \in L^2(\mathcal{D}) : v_{hp}|_\tau \circ \mathfrak{r} \in \mathcal{P}_p(\hat{\tau}),\ \tau \in \mathcal{T}_h\}.$$

Because of the insufficient smoothness of functions from $\mathcal{V}_{hp}^{dg}$, the Green's theorem used for obtaining the weak form must be applied element-wise, leading to the formulation (for a particular direction $\mathbf{\Omega} \in \omega$)

$$\sum_{\tau\in\mathcal{T}_h} \int_\tau (-u_{hp}\mathbf{\Omega}\cdot\nabla v_{hp} + \sigma_t u_{hp} v_{hp})\,\mathrm{d}\mathbf{r} + \sum_{e\not\subset\partial\mathcal{D}^-} \int_e \langle \mathbf{\Omega} u_{hp}\rangle \cdot [\![v_{hp}]\!]\,\mathrm{d}S = \int_\mathcal{D} q v_{hp}\,\mathrm{d}\mathbf{r}$$

where e denotes subsequently all faces of all elements $\tau \in \mathcal{T}_h$ (both interior and those coinciding with segments of $\partial\mathcal{D}$ [16]) and for each face

$$[\![v_{hp}]\!] = \begin{cases} v_{hp}\mathbf{n}^- + v_{hp}\mathbf{n}^+ & \text{for } e \not\subset \partial\mathcal{D}, \\ v_{hp}\mathbf{n} & \text{for } e \subset \partial\mathcal{D} \end{cases}$$

with $\mathbf{n}^\pm$ the outer normal of τ^+ and τ^-, respectively, where $e = \tau^- \cap \tau^+$. There are several ways of approximating $\mathbf{\Omega} u_{hp}$ by $\langle \mathbf{\Omega} u_{hp}\rangle$ (called *numerical flux*), the simplest stable approximation being the *upwind numerical flux*:

$$\langle \mathbf{\Omega} u_{hp}\rangle = \begin{cases} \mathbf{\Omega} u_{hp}^-, & \mathbf{\Omega}\cdot\mathbf{n}^- > 0, \\ \mathbf{\Omega} u_{hp}^+, & \mathbf{\Omega}\cdot\mathbf{n}^- < 0, \\ \mathbf{\Omega}\frac{u_{hp}^- + u_{hp}^+}{2}, & \mathbf{\Omega}\cdot\mathbf{n}^- = 0 \end{cases}$$

where $u_{hp}^\pm$ denotes the trace $u_{hp}|_e$ taken from τ^+ and τ^-, respectively. We refer e.g. to [56] for further details.

Another approach leaves the approximation space conforming (continuous), but modifies directly the bilinear form (typically by adding some artificial diffusion). This leads to stabilized continuous Galerkin methods (in which the bilinear form becomes dependent on mesh parameter h). We refer to [65] for the application of a particular conforming stabilized method – the streamline-upwind Petrov-Galerkin method – and further discussion about the discontinuous Galerkin method.

[16] we assume here vacuum conditions on $\partial\mathcal{D}^-$ for simplicity; using (3.61), reflective conditions could be incorporated straightforwardly

3.5.3 Diffusion approximation

Much simpler situation arises in the spatial discretization of the diffusion equation (3.28) with boundary conditions (3.30). In this case, the original problem is posed in the usual Sobolev space $H^1(\mathcal{D})$: Find $u \in H^1(\mathcal{D})$ such that

$$a(u,v) = f(v) \quad \forall v \in H^1(\mathcal{D}) \tag{3.67}$$

where

$$a(u,v) = (D\nabla u, \nabla v)_{L^2(\mathcal{D})} + (\Sigma u, v)_{L^2(\mathcal{D})} + (\gamma u, v)_{L^2(\partial\mathcal{D})}, \quad f(v) = (q,v)_{L^2(\mathcal{D})}.$$

(with $\Sigma = \sigma_t(\mathbf{r}) - \sigma_{s0}(\mathbf{r}) - \nu\sigma_f(\mathbf{r})$, D defined in eq. (3.28) and γ in (3.30)). Under the subcriticality conditions, the bilinear form a is bounded and coercive on $H^1(\mathcal{D})$ (even in the multigroup case – see [29, Chap. VII]). By using the finite element method as described above with the approximation space (3.64), an algebraic system

$$\mathbf{A}\mathbf{u} = \mathbf{b} \tag{3.68}$$

with sparse, symmetric, positive definite (in the mono-energetic case) matrix $\mathbf{A}$ is obtained, amenable to solution by standard numerical methods.

3.5.4 On the origin of errors in FE approximation

Let us finish this section by recalling a simple, yet very important connection between the above described finite element discretization and solution of (3.68). Even though the analysis is done here for the case of diffusion approximation with symmetric and positive definite system (3.68), keeping in mind its conclusion is equally important for finite-element discretizations of other models as well.

As before, after restricting to $\mathcal{V}_{hp} \subset H^1(\mathcal{D})$ we get the approximate problem: Find $u_{hp} \in \mathcal{V}_{hp}$ such that

$$a(u_{hp}, v_{hp}) = f(v_{hp}) \quad \forall v_{hp} \in \mathcal{V}_{hp}$$

(where $u_{hp} = \Pi_{hp} u$, $v_{hp} = \Pi_{hp} v$). Subtracting from (3.67), we obtain the well-known *Galerkin orthogonality* property:

$$a(u - u_{hp}, v_{hp}) = 0 \quad \forall v_{hp} \in \mathcal{V}_{hp},$$

which characterizes the *discretization error*. In practice, it is impossible to solve the system (3.68) exactly, so suppose that we have obtained after n steps of a suitable iterative method the solution $\mathbf{u}_{(n)}$, such that the *algebraic error* $\mathbf{u} - \mathbf{u}_{(n)}$ is nontrivial. By applying $\widehat{\mathcal{I}}_{hp}$ to the algebraic error, we obtain the representation of that error in $\mathcal{V}_{hp}$:

$$u_{hp} - u_{hp}^{(n)} \in \mathcal{V}_{hp}$$

(where we temporarily shifted the iteration index to improve readability). Hence, by decomposing the total error as

$$u - u_{hp}^{(n)} = (u - u_{hp}) + (u_{hp} - u_{hp}^{(n)})$$

and applying Galerkin orthogonality, we find that

$$a(u - u_{hp}^{(n)}) = a(u - u_{hp}) + a(u_{hp} - u_{hp}^{(n)}).$$

Also by noticing that

$$a(u_{hp} - u_{hp}^{(n)}) = \big(\mathbf{u} - \mathbf{u}_{(n)}\big)^T \mathbf{A} \big(\mathbf{u} - \mathbf{u}_{(n)}\big) = \|\mathbf{u} - \mathbf{u}_{(n)}\|_{\mathbf{A}}$$

(see (3.65)) we get the fundamental representation of the energy norm of the total error as a sum of the discretization error and algebraic error contributions:

$$a(u - u_{hp}^{(n)}) = a(u - u_{hp}) + \|\mathbf{u} - \mathbf{u}_{(n)}\|_{\mathbf{A}}.$$

In theory, the first part could be controlled by a suitable *hp*-adaptivity process (as e.g. in Sec. 6.2), while for the latter, using the methods that are based on minimization of the **A**-norm of error (the CG method among the Krylov subspace methods, or the smoothed aggregation multigrid method, as we demonstrate in App. F). However, striking the balance between the two contributions and determining optimal stopping criteria for the algebraic solution methods accordingly is an important area of active research (notably in the case when the latter is further split to account for rounding errors of computers) and we refer to paper [8] and its extensive list of references for further details.

4

The simplified P_N approximation

The simplified P_N (SP_N) approximation was proposed in the early 1960's by E. Gelbard [51, 52]) to circumvent the problem of increasing complexity of the P_N approximation in multiple dimensions. Its derivation was completely formal at the beginning – amounting to a simple replacement of differential operators $\frac{d}{dz}$ in the 1D P_N system by their multidimensional counterparts ∇ and $\nabla\cdot$ and recasting those scalar unknowns operated upon by the latter as vector quantities. Despite this mathematically weak derivation, the SP_N solution has been found to be equivalent to the solution of the multidimensional P_N equations in the case of a homogeneous medium and, comparing to either diffusion or P_N models, provided encouraging results both in terms of accuracy and efficiency even in more realistic cases. This is a rather remarkable fact – as we will see below, the SP_N approximation for odd N consists of $2N-1$ coupled elliptic partial differential equations (and reduces to the diffusion approximation for $N=1$), which is significantly lower than the $(N+1)^2$ equations of the full P_N model (and also than the $N(N+1)/2$ strongly coupled elliptic equations to which the full P_N model can be reduced). The method has thus become particularly attractive as its implementation required only modification of existing multigroup diffusion codes.

After some time, however, special transport problems for which the simple diffusion approximation actually provided better results have been contrived (see, e.g., [25, p. 247]). Validity of Gelbard's formal derivation therefore became questioned and the SP_N equations have not been seriously considered as a robust

enough improvement of the diffusion model for some time. This has changed in the 1990's with the extension of the asymptotic analysis originally performed for the diffusion approximation. Larsen, Morel and McGhee [42] have shown that under the scaling of the transport equation by a "diffusivity" parameter ε that makes the diffusion equation agree with the transport equation up to terms of order $\mathcal{O}(\varepsilon^3)$ as $\varepsilon \to 0$ (as already discussed in Sec. 3.4.3.1), the SP_3 equations are equivalent to the transport equation up to terms of order $\mathcal{O}(\varepsilon^7)$ provided that that the transport solution shows a nearly one-dimensional behaviour in the vicinity of interfaces of different materials by having there sufficiently weak tangential derivatives. The approach used by the authors was sufficiently general to show that SP_N equations of increasing order provide asymptotic corrections of the NTE of increasing order, which has been confirmed at least experimentally (e.g., [88] or [83]).

Brantley and Larsen [15] contributed to the theoretical justification of the method by variational analysis through which they showed that the SP_3 equations are the approximate Euler-Lagrange equations whose solution makes stationary a special physically reasonable functional characterizing arbitrary reaction rates. By including boundary terms in the functional, the authors also arrived at natural boundary conditions for the method, missing in the asymptotic approach. At internal interfaces, however, an assumption of one-dimensional behavior of solution was again required as in the asymptotic derivation of Larsen et al.

Together with other asymptotic and variational analyses (e.g. [93]), the range of validity of the approximation had been finally determined by the end of the 1990's. Although it turned out that this range is not significantly larger than that of the diffusion theory ([42]), the SP_N approximation has recently been shown to produce more accurate results than the diffusion model under these conditions and regained attention [48, 68, 73, 83, 88, 94].

The SP_N method is particularly suitable for solving reactor criticality problems, where the asymptotic conditions predominantly hold. Downar [38] compared the S_{16}, SP_3 and diffusion approximations over several model problems, with results that the SP_3 method well agrees with the high-order transport solution of the S_{16} method and provides more than 80% improvement in reactor

critical number and 50% to 30% improvement in pin powers[1] over the diffusion approximation. Somewhat smaller but still well-noticeable improvement has been obtained by Brantley and Larsen in [15]. Similarly to Downar and others, however, they conclude that SP_3 captures most of the transport effects in diffusive regimes of nuclear reactors (and that higher orders than 3 are not usually necessary, as also shown e.g. by Cho et al. in [22]). The authors also warned, however, that more careful spatial discretization than in the diffusion methods is required in order to capture the sharper boundary layer behaviour of the more transport-like SP_3 approximation.

Even though the SP_N solution does not tend to the exact solution of the NTE as $N \to \infty$ in general, there are several cases in which it is equivalent to the convergent P_N expansion (see some recent papers like [28, 36, 71]) and further research of the SP_N model and its connections to the NTE appears to be an interesting topic. One contribution of this work to this research is the description of a new way of deriving the SP_N equations from a specially formulated P_N approximation, which will be the subject of Chap. 5. In this chapter, we will recall the standard derivation and conclude by proving well-posedness of the weak form of the SP_N equations via the standard Lax-Milgram lemma.

4.1 Derivation of the SP_N equations

To illustrate the original Gelbard's approach, let us consider the case of one-dimensional symmetry, in which neutron transport is characterized by neutron distribution that is spatially varying only along one coordinate direction and, moreover, that is symmetric with respect to rotations about that axis. Without loss of generality, we may choose the principal direction of variation along the z-axis. This situation may arise for example when the system is composed of slabs, each with homogeneous properties and extents in the x and y directions much larger than in the principal direction, so that dependence on x and y may be neglected. We will identify the inflow/outflow boundaries $\partial\mathcal{D}^{\pm}$ with points $z_{\pm}$ and interior points $\mathbf{r} \in \mathcal{D}$ with $z \in (z_-, z_+)$.

[1] eq. (2.19) integrated over elementary cells comprising fuel assemblies

Under these assumptions, spherical harmonic functions reduce to Legendre polynomials in $\mu = \cos\vartheta = \Omega_z$ and partial derivatives $\frac{\partial}{\partial x}$, $\frac{\partial}{\partial y}$ vanish, so that the set of P_N equations (3.10) (where we assume N odd) becomes

$$\frac{n+1}{2n+1}\frac{d\phi_{n+1}(z)}{dz} + \frac{n}{2n+1}\frac{d\phi_{n-1}(z)}{dz} + \Sigma_n(z)\phi_n(z) = q_n(z), \tag{4.1}$$

where $n = 0, 1, \ldots, N$ (discarding the non-sensical moments ϕ_n for negative n),

$$\Sigma_n = \sigma_t - \kappa_n = \sigma_t - \sigma_{sn} - \delta_{n0}\nu\sigma_f$$

and the moments are defined as

$$\phi_n = \int_{-1}^{1} P_n(\mu)\psi(\cdot,\mu)\mathrm{d}\mu\,, \quad q_n = \int_{-1}^{1} P_n(\mu)q(\cdot,\mu)\mathrm{d}\mu$$

(note that the definition of κ_n by (3.20) is still valid with $\mu_0 = \mu\mu'$).

To proceed as in the derivation of the diffusion equation, we again assume that $\Sigma_n \geq \underline{\Sigma_n} > 0$ for $n = 1, 3, \ldots, N$ and $q_n = 0$ for $n \geq 1$. Then by solving the odd-order equations for the odd-order flux moments in terms of a derivative of the even-order flux moments and using the result to eliminate the odd-order flux moments from the even-order equations, we obtain the one dimensional SP_N equations. To write them in a convenient form, we define the auxiliary SP_N moments:

$$\phi_n^s := (n+1)\phi_n + (n+2)\phi_{n+2}, \quad n = 0, 2, \ldots, N-1, \tag{4.2}$$

(setting $\phi_{N+1} = 0$) and SP_N "diffusion coefficients"

$$D_n^s := \frac{1}{(2n+1)\Sigma_n}, \quad n = 1, 3, \ldots, N \tag{4.3}$$

so that SP_N currents could be defined as

$$J_n^s \equiv \phi_{2n+1} = -D_{2n+1}^s \frac{d\phi_n^s}{dz}, \quad n = 0, 2, \ldots, N-1. \tag{4.4}$$

Notice that $J_1^s = J$ (neutron current in the one-dimensional symmetry) and that scalar flux is given by

$$\phi \equiv \phi_0 = \phi_0^s - \frac{2}{3}\phi_2^s + \frac{8}{15}\phi_4^s - \ldots = \sum_{n=0}^{(N-1)/2} F_n\phi_{2n}^s, \quad F_n = (-1)^n\frac{2^n n!}{(2n+1)!!} \tag{4.5}$$

where N is odd and $(2n+1)!! = (2n+1)(2n-1)\cdots 3\cdot 1$.

One may note that these definitions are somewhat arbitrary and indeed, there have been several "SP_N approximations" reported in literature. We have compared several of them and found that they are all equivalent[2]. The formulation obtained with the above definitions is particularly convenient as it allows to easily obtain well-posedness of its corresponding weak form (at least under some additional constraints on the higher-order anisotropic scattering moments).

4.1.1 The SP_3 case

As the practical usefulness of the SP_N equations has been experimentally verified to be limited by orders up to around $N = 7$ (we refer to above mentioned papers and reports), we do not delve into technical derivation of the general form of the SP_N equations here and rather consider the case $N = 3$ (with cases $N = 5, 7$ included in App. C [3]).

The one-dimensional P_3 system reads

$$\begin{aligned}
\frac{d\phi_1(z)}{dz} + \Sigma_0(z)\phi_0(z) &= q_0(z),\\
\frac{1}{3}\frac{d\phi_0(z)}{dz} + \frac{2}{3}\frac{d\phi_2(z)}{dz} + \Sigma_1(z)\phi_1(z) &= q_1(z),\\
\frac{2}{5}\frac{d\phi_1(z)}{dz} + \frac{3}{5}\frac{d\phi_3(z)}{dz} + \Sigma_2(z)\phi_2(z) &= q_2(z),\\
\frac{3}{7}\frac{d\phi_2(z)}{dz} + \Sigma_3(z)\phi_3(z) &= q_3(z)
\end{aligned} \tag{4.6}$$

and the Marshak approximation of albedo boundary conditions (2.8)

$$\begin{aligned}
\frac{\phi_0(z_\pm)}{4} + \frac{5\phi_2(z_\pm)}{16} \mp \frac{\phi_1(z_\pm)}{2} &= \beta(z_\pm)\left[\frac{\phi_0(z_\pm)}{4} \pm \frac{\phi_1(z_\pm)}{2} + \frac{5\phi_2(z_\pm)}{16}\right]\\
-\frac{\phi_0(z_\pm)}{16} + \frac{5\phi_2(z_\pm)}{16} \mp \frac{\phi_3(z_\pm)}{2} &= \beta(z_\pm)\left[-\frac{\phi_0(z_\pm)}{16} + \frac{5\phi_2(z_\pm)}{16} \pm \frac{\phi_3(z_\pm)}{2}\right].
\end{aligned} \tag{4.7}$$

[2] For instance, going from "our" SP_3 system to that used by Brantley [15] for his variational analyses (which is actually the same as that used by Larsen, Morel and McGhee in [42] for asymptotic analyses) amounts to multiplying the first SP_3 equation of Brantley by $5/9$ and the second by 3 and the use of (4.2) (and analogously for the boundary conditions).

[3] The equations were generated by a simple Mathematica script that can be used for any reasonably high order N if needed.

Using the approach described above together with the auxiliary SP_3 definitions, we obtain the following one-dimensional SP_3 system

$$\begin{aligned} -\frac{d}{dz}\mathbf{D}^s(z)\frac{d}{dz}\Phi^s(z) + \mathbf{C}^s(z)\Phi^s(z) &= \mathrm{Q}^s(z), \quad z \in (z_-, z_+) \\ \mathbf{D}^s(z)\frac{d}{dz}\Phi^s(z) + \gamma(z_\pm)\mathbf{G}^s\Phi^s(z_\pm) &= \mathbf{0} \quad \gamma(z\pm) = \frac{1-\beta(z_\pm)}{2(1+\beta(z_\pm))} \end{aligned} \tag{4.8}$$

where γ is the same albedo coefficient as in the diffusion case (3.30) and

$$\begin{aligned} \Phi^s = [\phi_0^s, \phi_2^s]^T, \quad \mathrm{Q}^s = [q_0, -\tfrac{2}{3}q_0]^T, \quad \mathbf{D}^s = \mathrm{diag}\left\{\frac{1}{3\Sigma_1}, \frac{1}{7\Sigma_3}\right\}, \\ \mathbf{C}^s = \begin{bmatrix} \Sigma_0 & -\frac{2\Sigma_0}{3} \\ -\frac{2\Sigma_0}{3} & \frac{4\Sigma_0}{9} + \frac{5\Sigma_2}{9} \end{bmatrix}, \quad \mathbf{G}^s = \begin{bmatrix} 1 & -\frac{1}{4} \\ -\frac{1}{4} & \frac{7}{12} \end{bmatrix}. \end{aligned} \tag{4.9}$$

Using the Gelbard's ad-hoc approach, the multidimensional equations are just

$$\begin{aligned} -\nabla\cdot\mathbf{D}^s(\mathbf{r})\nabla\Phi^s(\mathbf{r}) + \mathbf{C}^s(\mathbf{r})\Phi^s(\mathbf{r}) &= \mathrm{Q}^s(\mathbf{r}), \quad \mathbf{r}\in\mathcal{D}, \\ \mathbf{n}(\mathbf{r})\cdot\mathbf{D}^s(\mathbf{r})\nabla\Phi^s(\mathbf{r}) + \gamma(\mathbf{r})\mathbf{G}^s\Phi^s(\mathbf{r}) &= \mathbf{0}, \quad \mathbf{r}\in\partial\mathcal{D}, \end{aligned} \tag{4.10}$$

where $\nabla\Phi^s$ is the Jacobian matrix of Φ^s:

$$[\nabla\Phi^s]_{i,\alpha} = \frac{\partial\phi^s_{2i-2}}{\partial x_\alpha}, \quad i = 1,2, \quad \alpha = 1,2,3$$

(we start using the convention that Greek subscripts index the Cartesian coordinate axes, which will become more copious in Chap. 5),

$$\nabla = \left[\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right], \ \mathbf{n} = [n_x, n_y, n_z]$$

and for $\mathbf{v} = [v_x, v_y, v_z]$

$$\mathbf{v}\cdot\mathbf{A} = \sum_{\alpha=1}^{3} v_\alpha A_{i\alpha}.$$

4.2 Weak formulation

Let $\mathbb{H}^1(\mathcal{D}) = [H^1(\mathcal{D})]^2$. To introduce some new notation, let us write the inner product on this space as

$$(\mathrm{U}, \mathrm{V})_{\mathbb{H}^1(\mathcal{D})} = \int_{\mathcal{D}} (\nabla \mathrm{U} : \nabla \mathrm{V} + U \cdot V) \, \mathrm{d}\mathbf{r} \,; \tag{4.11}$$

here $\cdot$ denotes the usual inner product and : the double inner product of matrices:

$$\mathbf{A} : \mathbf{B} = \sum_{i,j} A_{ij} B_{ij}. \tag{4.12}$$

The weak formulation forming the basis for the finite element solution can now be stated as follows:

Problem 4. *Given $q_0 \in L^2(\mathcal{D})$, find $\Phi^s = \mathrm{col}\,\{\phi_0^s, \phi_2^s\} \in \mathbb{H}^1(\mathcal{D})$ such that*

$$\begin{gathered} a(\Phi^s, \mathrm{V}) = f(\mathrm{V}) \quad \forall \mathrm{V} \in \mathbb{H}^1(\mathcal{D}), \\ a(\mathrm{U}, \mathrm{V}) := \int_{\mathcal{D}} \left(\mathbf{D}^s \nabla \mathrm{U} : \nabla \mathrm{V} + \mathbf{C}^s \mathrm{U} \cdot \mathrm{V} \right) d\mathbf{r} \, + \int_{\partial \mathcal{D}} \gamma \mathbf{G}^s \mathrm{U} \cdot \mathrm{V} \, dS \,, \\ f(\mathrm{V}) := \int_{\mathcal{D}} \mathrm{Q}^s \cdot \mathrm{V} \, d\mathbf{r} \,, \quad \left(\mathbf{D}^s \nabla \mathrm{U} \right) : \nabla \mathrm{V} = \sum_{i=1}^{2} \sum_{\alpha=1}^{3} D_{2i-1}^s \frac{\partial u_i}{\partial x_\alpha} \frac{\partial v_i}{\partial x_\alpha}. \end{gathered} \tag{4.13}$$

Note that eq. (4.13) represent a set of weakly coupled diffusion-like equations. The case of $N = 1$ also reduces to the weak form of the usual diffusion approximation, eq. (3.67). We also note that in the multigroup approximation of energetic dependence, the diffusion approximation has the same form (4.10) (with weak form (4.13)), with $\mathbf{D}^s$, $\mathbf{C}^s$, $\gamma\mathbf{G}^s$, Q^s and Φ^s replaced by

$$\begin{gathered} \mathbf{D} = \mathrm{diag}\,\{D^g\}_G, \quad [\mathbf{C}]_{gg'} = \sigma_t^g \delta_{gg'} - \sigma_s^{gg'} - \chi^g \nu \sigma_f^{g'}, \quad [\gamma\mathbf{G}]_{gg'} = \gamma^{gg'}, \\ \mathrm{Q} = \mathrm{col}\,\{q_0^g\}_G \,, \quad \Phi = \mathrm{col}\,\{\phi^g\}_G \,, \end{gathered} \tag{4.14}$$

where $g, g' = 1, 2, \ldots, G$. The extension to the multigroup SP_N case is obvious, with the weak formulation posed in $\mathbb{H}^1(\mathcal{D}) = [H^1(\mathcal{D})]^{(2N-1)\times G}$.

Also note that Problem 4 is equivalent to

Problem 4′. *Given* $q_0 \in L^2(\mathcal{D})$, *find* $\Phi^s = \mathrm{col}\,\{\phi_0^s, \phi_2^s\} \in \mathbb{H}^1(\mathcal{D})$ *such that*

$$\begin{aligned} a_{00}(\phi_0^s, \varphi_0) + a_{02}(\phi_2^s, \varphi_0) &= f_0(\varphi_0), \quad \forall \varphi_0 \in H^1(\mathcal{D}), \\ a_{20}(\phi_0^s, \varphi_2) + a_{22}(\phi_2^s, \varphi_2) &= f_2(\varphi_2), \quad \forall \varphi_2 \in H^1(\mathcal{D})\,^4, \\ a_{ij}(u,v) := \int_{\mathcal{D}} \left(D_{ij}^s \nabla u \nabla v + C_{ij}^s uv\right) d\mathbf{r} &+ G_{ij}^s \int_{\partial\mathcal{D}} \gamma uv \, dS\,, \\ f_i(v) := \int_{\mathcal{D}} q_i^s v \, d\mathbf{r}\,, &\quad i = 0, 2,\ j = 0, 2. \end{aligned} \tag{4.15}$$

To see this, if Φ^s is the solution of Problem 4, then for $V = [\varphi_0, 0]^T$ and $V = [0, \varphi_2]^T$, respectively, we satisfy the first and second equation of Problem 4′, respectively. On the other hand, if Φ^s is the solution of Problem 4′, then summing up the equations (4.15) shows that Φ^s also solves Problem 4. The same approach can be obviously used to generate appropriate formulations of the higher-N or multigroup problems.

When using the finite element method to obtain approximate solution of Problem 4 (or equivalently 4′), the formulation is restricted to a finite-dimensional subspace of $\mathbb{H}^1(\mathcal{D})$ as discussed in Sec. 3.5.2 and an element-by-element assembling procedure is employed to obtain a system of discrete algebraic equations. Two finite element frameworks have been considered in this thesis for this task – Hermes2D (which the author helped developing and uses for testing *hp*-adaptivity ideas) and Dolfin (which, thanks to its 3D support, has been used as a basis for the research project referred to in the Introduction). Problem 4′ is in the form suited for implementation in the Hermes2D framework, while the form of Problem 4 is appropriate for the Dolfin framework (Chap. 6).

4.3 Well-posedness of the SP_3 formulation

In this section, we will study the properties of coupling matrices of the (mono-energetic) SP_3 method, which allow us to establish well-posedness of Problem

[4] The reason for the here superfluous distinction between the test functions φ_0 and φ_2 becomes clear when the approximate problem is formulated on finite-dimensional subspaces of $H^1(\mathcal{D})$; different approximation order can then be used for approximating the zero-th and second order moments and we utilize this fact in Chap. 6.

4. We could not find a formal proof of this property in the available literature, which motivated this study concluded in Theorem 6 and Remark 11 (discussing the extension to higher SP_N orders) at the end of this section.

The SP_3 matrices $\mathbf{D}^s$, $\mathbf{C}^s$ and $\mathbf{G}^s$ are symmetric, with elements bounded a.e. in $\mathcal{D}$, hence the bilinear form $a(\mathrm{U},\mathrm{V})$ is also bounded on $\mathbb{H}^1(\mathcal{D})$ (with norm induced by the inner product (4.11); for the boundary term, recall that $\gamma \in [0, 0.5]$ with value 0 at perfectly reflecting boundary and value 0.5 at vacuum boundary and use the standard trace inequality in $H^1(\mathcal{D})$ on each term in the sum of boundary integrals). The linear form f is also obviously bounded when the isotropic source term $q_0 \in L^2(\mathcal{D})$.

For coercivity, note that the matrix $\mathbf{D}^s$ is positive definite, as is $\mathbf{G}^s$ (being symmetric and strictly diagonally dominant, positive-definiteness follows from the Gerschgorin theorem).

To show positive-definiteness of $\mathbf{C}^s$, let us again split

$$\Sigma_n = \sigma_t - \kappa_n.$$

Then

$$\mathbf{C}^s = \sigma_t \mathbf{C}_t^s - \kappa_0 \mathbf{C}_0^s - \kappa_2 \mathbf{C}_2^s, \tag{4.16}$$

where

$$\begin{aligned} \mathbf{C}_0^s &= \begin{bmatrix} 1 & -\frac{2}{3} \\ -\frac{2}{3} & \frac{4}{9} \end{bmatrix}, \quad \mathbf{C}_2^s = \begin{bmatrix} 0 & 0 \\ 0 & \frac{5}{9} \end{bmatrix}, \\ \mathbf{C}_t^s &= \mathbf{C}_0^s + \mathbf{C}_2^s = \begin{bmatrix} 1 & -\frac{2}{3} \\ -\frac{2}{3} & 1 \end{bmatrix}. \end{aligned} \tag{4.17}$$

Using the Cauchy-Schwarz inequality and the addition theorem (A.5), we obtain

$$\begin{aligned} \frac{2n+1}{4\pi} |P_n(\mu_0)| &= \left| \sum_{m=-n}^{n} Y_n^m(\boldsymbol{\Omega}) Y_n^m(\boldsymbol{\Omega}') \right| \\ &\leq \sqrt{\sum_{m=-n}^{n} \left[Y_n^m(\boldsymbol{\Omega}) \right]^2} \sqrt{\sum_{m=-n}^{n} \left[Y_n^m(\boldsymbol{\Omega}') \right]^2} \\ &= \frac{2n+1}{4\pi} \sqrt{P_n(\boldsymbol{\Omega}\cdot\boldsymbol{\Omega}) P_n(\boldsymbol{\Omega}'\cdot\boldsymbol{\Omega}')} \\ &= \frac{2n+1}{4\pi} |P_n(\boldsymbol{\Omega}\cdot\boldsymbol{\Omega})| \\ &= \frac{2n+1}{4\pi} P_n(1) \end{aligned}$$

with equality occurring in the case $\mu_0 = \mathbf{\Omega} \cdot \mathbf{\Omega}' = 1$. It follows that

$$P_n(\mu_0) \leq |P_n(\mu_0)| < P_n(1) = 1 = P_0(\mu_0), \quad n \geq 1,\ \mu_0 \in (-1, 1).$$

Since the fission part of the collision kernel κ is isotropic, we thus have for $n \geq 1$

$$\kappa_n = \sigma_{sn} = 2\pi \int_{-1}^{1} \sigma_s(\mu_0) P_n(\mu_0) \mathrm{d}\mu_0 \ < 2\pi \int_{-1}^{1} \sigma_s(\mu_0) \mathrm{d}\mu_0 \ = \sigma_{s0} < \sigma_{s0} + \nu\sigma_f = \kappa_0. \tag{4.18}$$

If we now define the relation "$<$" between matrices as

$$\mathbf{A} < \mathbf{B} \quad \Leftrightarrow \quad \mathbf{x}^T \mathbf{A} \mathbf{x} < \mathbf{x}^T \mathbf{B} \mathbf{x}, \ \ \forall \mathbf{x} \neq \mathbf{0},$$

we have from (4.18), (4.16) and (4.17)

$$\mathbf{C}^s > \sigma_t \mathbf{C}_t^s - \kappa_0 (\mathbf{C}_0^s + \mathbf{C}_2^s) = (\sigma_t - \kappa_0) \mathbf{C}_t^s.$$

Under the subcriticality conditions in $L^2(X_{|E})$ (Def. 1), we have

$$\sigma_t > \sigma_s + \nu\sigma_f = \kappa_0$$

(cf. the representation (2.24)). As the matrix $\mathbf{C}_t^s > 0$ (again because its strict diagonal dominance, symmetry and positivity of diagonal elements), we have thus proved that $\mathbf{C}^s$ is also positive definite. From the Lax-Milgram lemma (Lemma 1 on pg. 23), we directly obtain the following theorem.

Theorem 6. *Let the subcriticality conditions in $L^2(X_{|E})$ hold. Then Problem 4 has a unique solution and there exists constant $\alpha > 0$ such that*

$$\|\Phi^s\|_{\mathbb{H}^1(\mathcal{D})} \leq \frac{1}{\alpha} \|\mathrm{Q}^s\|_{\mathbb{L}^2(\mathcal{D})}.$$

Remark 11. We note that for higher order SP_N approximations, the previously stated properties for $\mathbf{D}^s$, $\mathbf{G}^s$ are still valid and the decomposition (4.16) has the following form:

$$\mathbf{C}^s = \sigma_t \mathbf{C}_t^s - \kappa_0 \mathbf{C}_0^s - \kappa_2 \mathbf{C}_2^s - \kappa_4 \mathbf{C}_4^s \ldots - \kappa_{2N-1} \mathbf{C}_{2N-1}^s.$$

The matrix

$$\mathbf{C}_t^s = \sum_{n=0}^{\lfloor N/2 \rfloor} \mathbf{C}_{2n}^s$$

(where $\lfloor N/2 \rfloor$ is the integer part of $N/2$) is no longer strictly diagonally dominant, but it is still positive definite, as can be verified by explicitly computing the eigenvalues (see App. C for the cases $N = 5, 7$; the Mathematica script referenced in the appendix may be easily used for higher N).

5

The MCP_N approximation

The main goal of this chapter is to derive an alternative form of the P_N approximation that provides additional insight into the structure of the equations. In particular, it allows to derive the SP_3 equations in a new way that we will present at the end of the chapter. As in the exposition of the original P_N method, we will focus on the monoenergetic equation $T\psi - q = 0$ a.e. in $X_{|E}$, that is

$$\begin{aligned}&\boldsymbol{\Omega}\cdot\nabla\psi(\mathbf{r},\boldsymbol{\Omega}) + \sigma_t(\mathbf{r})\psi(\mathbf{r},\boldsymbol{\Omega})\\ &\quad - \int_{\mathcal{S}_2}\left[\sigma_s(\mathbf{r},\boldsymbol{\Omega}\cdot\boldsymbol{\Omega}') + \frac{\nu\sigma_f(\mathbf{r})}{4\pi}\right]\psi(\mathbf{r},\boldsymbol{\Omega}')\,\mathrm{d}\boldsymbol{\Omega}' - q(\mathbf{r},\boldsymbol{\Omega}) = 0, \quad (\mathbf{r},\boldsymbol{\Omega})\in\mathcal{D}\times\mathcal{S}_2\end{aligned} \tag{5.1}$$

and assume that $\psi(\mathbf{r},\cdot)\in L^2(\mathcal{S}_2)$ (we have split the collision kernel κ into the scattering and fission part so that we can show how non-isotropic and isotropic terms are handled).

Notation conventions

We introduce some new notation in this chapter. The new rules added to the notation conventions of Chap. 3 are summarized in the following list.

- $\mathbb{A}^{(n)}$ ... Cartesian tensor of rank n (Def. 4 below),
- $A^{(n)}_{\alpha_1\cdots\alpha_n}$ ... Cartesian tensor $\mathbb{A}^{(n)}$ in an indexed notation,
- $\underset{m}{\cdot}$ ... m-fold contraction (generalization of inner product, Def. 7),
- $\mathscr{S}\left(\mathbb{A}^{(n)}\right)$ ... symmetrization of tensor $\mathbb{A}^{(n)}$ (Def. 12),

- tr $\mathbb{A}^{(n)}$ ... trace of tensor $\mathbb{A}^{(n)}$ (Def. 9),
- Greek subscripts (as in $A^{(1)}_\alpha$) ... indices of the Cartesian axes (with the correspondence $\alpha = 1, 2, 3 \leftrightarrow x, y, z$).

Let us start by recalling basic ingredients of the spherical harmonics method from Sec. 3.3, relevant for subsequent sections.

5.1 Classical P$_N$ approximation

Being a square integrable function on $L^2(\mathcal{S}_2)$, $\psi(\mathbf{r}, \cdot)$ may be represented by a generalized Fourier series expansion in terms of a complete orthonormal basis of this space. The P$_N$ method uses the basis of (tesseral) spherical harmonics $\{Y_n^m\}$ of degree $n \geq 0$ and order $-n \leq m \leq n$. A (semi-)finite approximation is obtained by considering the NTE in a subspace $L^2_K(\mathcal{S}_2) \subset L^2(\mathcal{S}_2)$ spanned by $\{Y_n^m\}$, $n \leq N$ (where we assume the more common case of odd N and $K = (N+1)^2$ in general three dimensional setting) – i.e. it consists of the expansion

$$\psi(\mathbf{r}, \mathbf{\Omega}) \approx \sum_{n=0}^{N} \sum_{m=-n}^{n} \phi_n^m(\mathbf{r}) Y_n^m(\mathbf{\Omega}) \tag{5.2}$$

and orthogonal projection of the residual $T\psi - q$ onto $L^2_K(\mathcal{S}_2)$ (similarly, the exact boundary conditions are projected onto $L^2_K(\partial X^-)$, orthogonally to the subspace spanned by the complete set of even-degree spherical harmonics).

We recall the *addition theorem* for Legendre polynomials P_n:

$$P_n(\mathbf{\Omega} \cdot \mathbf{\Omega}') = \frac{4\pi}{2n+1} \sum_{m=-n}^{n} Y_n^m(\mathbf{\Omega}) Y_n^m(\mathbf{\Omega}'), \tag{5.3}$$

which allows to simplify the collision integral, after expansion of its (non-isotropic) kernel in terms of Legendre polynomials up to the degree of scattering anisotropy N_s:

$$\sigma_s(\mathbf{r}, \mathbf{\Omega} \cdot \mathbf{\Omega}') \approx \sum_{n=0}^{N_s} \frac{2n+1}{4\pi} \sigma_{sn}(\mathbf{r}) P_n(\mathbf{\Omega}' \cdot \mathbf{\Omega}) \tag{5.4}$$

Also, we have the following relation between the spherical harmonic moments

$$\phi_n^m(\mathbf{r}) = \int_{\mathcal{S}_2} \psi(\mathbf{r}, \mathbf{\Omega}) Y_n^m(\mathbf{\Omega}) \, \mathrm{d}\mathbf{\Omega}$$

(unknowns in the P_N equations) and the important physical quantities (scalar flux, neutron current):

$$\phi = \sqrt{4\pi}\phi_0^0, \quad \mathrm{J} = \sqrt{\frac{4\pi}{3}} \begin{bmatrix} \phi_1^1 \\ \phi_1^{-1} \\ \phi_1^0 \end{bmatrix}.$$

5.2 Tensor form of spherical harmonics

This standard procedure results in a system of P_N equations given by eq. (3.10). It is possible to reduce this system of first-order PDEs into a system of second-order equations, which gets however quite complicated ([19, 86]) and in no way resembles the simple elliptic system of the SP_N approximation. This motivates the search for an alternative form of the expansion (5.2) that would reveal some connection with the computationally attractive SP_N set.

5.2.1 Surface and solid spherical harmonics

A possible way of achieving this goal starts by studying the linear combinations of spherical harmonics of fixed degree. We will simplify the notation by setting $\mu = \cos\vartheta$, replace $\mathbf{\Omega}$ with (μ, φ) where necessary and also use the following relations between the velocity and direction vectors in Cartesian coordinates:

$$\mathbf{v} = [v_x, v_y, v_z]^T = v\mathbf{\Omega}, \quad \mathbf{\Omega} = [\Omega_x, \Omega_y, \Omega_z]^T = \frac{\mathbf{v}}{v}.$$

Definition 3. [49, Def. 3.22] A general linear combination of the $2n+1$ (tesseral) spherical harmonics of degree n is called a *surface spherical harmonic of degree* n and can be written as

$$\mathcal{Y}_n(\mu, \varphi) = A_0 P_n(\mu) + \sum_{m=1}^{n} \left[A_m \cos(m\varphi) P_n^m(\mu) + B_m \sin(m\varphi) P_n^m(\mu) \right] \tag{5.5}$$

where $A_m, B_m \in \mathbb{R}$ and P_n^m are the *associated Legendre polynomials* (eq. (A.2)).

Note that there is a one-to-one relationship between the coefficients in eq. (5.2) for fixed n and the coefficients in (5.5).

Multiplying by v^n, we obtain the (regular) *solid spherical harmonic*[1]. Solid spherical harmonics were utilized in the early works [33, 99] as tools for analyzing the multidimensional spherical harmonics. Since then, it appears there has been no interest in solid spherical harmonics for approximating angular dependence of the NTE, until the paper by Ackroyd [1]. Ackroyd used these functions to arrive at a set of equations that can be used for practical approximation of the solution of the NTE. Under the assumption of isotropic scattering ($\sigma_{sk} = 0$ for $k \geq 1$ in (5.4)) Ackroyd derived for a homogeneous region a set of coupled diffusion-like equations (called SHP$_N$) without any other requirement (unlike the classical derivation of the SP$_N$ equations, which required certain assumptions about dimensionality or material properties, cf. the overview at the beginning of Sec. 4), together with heuristic boundary and interface conditions. As was shown in the paper, the SHP$_N$ equations reduced by simple substitutions to the set of SP$_N$ equations originally formulated by Gelbard and it is interesting to note that this actually showed that the latter are within an isotropically scattering homogeneous medium equivalent to the full solid harmonics expansion – the same result for surface spherical harmonics has been independently proved in [25] and revisited recently ([28, 83]). Unfortunately, the general treatment in [1] is very technically involved and in author's opinion quite difficult to follow – this may be the reason why the idea has not been picked up and possible research directions outlined in the paper's conclusion not pursued.

5.2.2 Cartesian tensors

In order to describe a conceptually simpler and arguably also more useful approach, we need to recall some basic facts about Cartesian tensors. We recall the convention that Greek subscripts (ranging from 1 to 3) represent axes of the Cartesian coordinate system with unit vectors $\mathbf{e}_x$, $\mathbf{e}_y$, $\mathbf{e}_z$. We will also use the

[1]Regular solid spherical harmonics are one class of solutions of the Laplace equation $\nabla^2 Y = 0$ in spherical coordinates (v, φ, ϑ) which vanish as $v \to 0$. The other are the irregular solid spherical harmonics, which have singularity of the form v^{-n-1} at the origin ([17, Chap. VI]).

convenient Einstein's summation convention which implies summation over any index that appears twice in an indexed expression, for instance

$$A^{(3)}_{\alpha\beta\gamma}B^{(2)}_{\beta\gamma} = \sum_{\beta=1}^{3}\sum_{\gamma=1}^{3} A^{(3)}_{\alpha\beta\gamma}B^{(2)}_{\beta\gamma} = C^{(1)}_{\alpha}.$$

Definition 4. An n-dimensional array $\mathbb{A}^{(n)}$ of 3^n components $A^{(n)}_{\alpha_1\ldots\alpha_n}$ is called *Cartesian tensor of rank n* if it transforms as

$$A^{(n)'}_{\alpha_1\ldots\alpha_n} = g_{\alpha_1\beta_1}\cdots g_{\alpha_n\beta_n}A^{(n)}_{\beta_1\ldots\beta_n} \tag{5.6}$$

under the change of coordinate system $Oxyz \to Ox'y'z'$ by the action of an orthogonal matrix $\mathbf{G} = [g_{\alpha\beta}]$:

$$\mathbf{e}'_\alpha = g_{\alpha\beta}\mathbf{e}_\beta.$$

A special case of the matrix $\mathbf{G}$ was the rotation matrix $\mathbf{R}$ introduced in Sec. 2.2.7; here we include also the reflections about origin to make the definition general. As we will only use the Cartesian tensors, we will henceforth omit the word Cartesian. Word "tensor" will also be used for a general *tensor field*, components of which are functions – like $\mathbb{A}^{(n)}(\mathbf{r})$. The transformation has then the following form:

$$A^{(n)'}_{\alpha_1\ldots\alpha_n}(\mathbf{r}) = A^{(n)}_{\alpha_1\ldots\alpha_n}(\mathbf{G}^T\mathbf{r}) = g_{\alpha_1\beta_1}\cdots g_{\alpha_n\beta_n}A^{(n)}_{\beta_1\ldots\beta_n}(\mathbf{r}).$$

We will denote by $\mathbb{I} \equiv \mathbb{I}^{(2)}$ the identity rank-2 tensor (matrix) and by $\mathbb{O}^{(n)}$ the zero rank-n tensor. The identity tensor can be written in the component notation via the Kronecker delta symbol:

$$I_{ij} = \delta_{ij}.$$

Addition and subtraction of two tensors of same rank and multiplication of a tensor by a scalar are done component-wise. Multiplication of two tensors is defined as follows:

Definition 5. Components of tensor $\mathbb{C}^{(n+m)} = \mathbb{A}^{(n)} \otimes \mathbb{B}^{(m)}$ are given by

$$C^{(n+m)}_{\alpha_1\ldots\alpha_n\beta_1\ldots\beta_m} = A^{(n)}_{\alpha_1\ldots\alpha_n}B^{(n)}_{\beta_1\ldots\beta_m}.$$

Definition 6. m-th power of tensor $\mathbb{A}^{(n)}$ is a tensor of rank nm defined as

$$\mathbb{C}^{(nm)} = \mathbb{A}^{(n)} \otimes \mathbb{A}^{(n)} \otimes \cdots \otimes \mathbb{A}^{(n)} \quad (m\text{-times})$$

We will mainly use powers of vectors and also consider the gradient operator as a vector

$$\nabla = \left[\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right]^T;$$

hence the Laplacian

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} = \nabla \cdot \nabla,$$

while the Hessian operator (second rank tensor operator) with components

$$D^{(2)}_{\alpha\beta} = \frac{\partial^2}{\partial x_\alpha \partial x_\beta} = \nabla \otimes \nabla.$$

Also, for simplicity, $\nabla\mathbb{A}^{(n)} \equiv \nabla \otimes \mathbb{A}^{(n)}$.

Definition 7. For $1 \leq m \leq n$, the m-fold contraction of tensors $\mathbb{A}^{(n)}$ and $\mathbb{B}^{(n)}$ is a rank-$(2n-2m)$ tensor $\mathbb{C}^{(2n-2m)} = \mathbb{A}^{(n)} \underset{m}{\cdot} \mathbb{B}^{(n)}$ with components

$$C^{(2n-2m)}_{\alpha_1 \ldots \alpha_{n-m} \beta_1 \ldots \beta_{n-m}} = A^{(n)}_{\alpha_1 \ldots \alpha_{n-m} \gamma_{n-m+1} \ldots \gamma_n} B^{(n)}_{\gamma_n \ldots \gamma_{n-m+1} \beta_{n-m} \ldots \beta_1}.$$

Specially for $n = m = 1$, we get the standard inner product of vectors (denoted by $\cdot$ so far), while for $n = m = 2$ we get the double inner product (denoted by :) of matrices (4.12). Generally when $n = m$, we obtain the scalar $A^{(n)}_{\gamma_1 \ldots \gamma_n} B^{(n)}_{\gamma_n \ldots \gamma_1}$ and suppress the index under the $\cdot$ sign to simplify the writing (it should be clear from the two operands and their rank that a total contraction over all their indices is intended).

Definition 8. For $m \leq \lfloor n/2 \rfloor$ (the integer part of $n/2$), the m-fold contraction of tensor $\mathbb{A}^{(n)}$ (contraction in m index pairs) is a rank-$(n-2m)$ tensor with components

$$B^{(n-2m)}_{\alpha_{2m+1} \ldots \alpha_n} = A^{(n)}_{\alpha_1 \alpha_1 \ldots \alpha_m \alpha_m \alpha_{2m+1} \ldots \alpha_n}.$$

When contracting in only 1 index pair, the 1-fold contraction is simply called a contraction. The result has a special name:

Definition 9. Contraction of tensor $\mathbb{A}^{(n)}$ in one index pair is called *trace* of $\mathbb{A}^{(n)}$ in that pair :

$$\text{tr } \mathbb{A}^{(n)} = A^{(n)}_{\alpha_1\alpha_2\alpha_3\ldots\alpha_n}\delta_{\alpha_1\alpha_2} = A^{(n-2)}_{\alpha_1\alpha_1\alpha_3\ldots\alpha_n}.$$

If the trace of $\mathbb{A}^{(n)}$ in one index pair vanishes, the tensor is said to be traceless in that pair. Obviously then, if the tensor does not depend on the order of indices, it is traceless in all index pairs.

Definition 10. Tensor $\mathbb{A}^{(n)}$ is *totally symmetric* if it is invariant under any permutation $\pi(\alpha_1 \ldots \alpha_n)$ of its indices:

$$A^{(n)}_{\alpha_1\ldots\alpha_n} = A^{(n)}_{\pi(\alpha_1\ldots\alpha_n)}.$$

Definition 11. Totally symmetric tensor whose trace in any index pair vanishes is called *totally symmetric traceless tensor* and its trace vanishes in all index pairs. We will denote totally symmetric traceless tensors as TST tensors.

We can make any tensor of rank n totally symmetric by applying the symmetrization operator $\mathscr{S}(\cdot)$:

Definition 12. $\widetilde{\mathbb{A}}^{(n)} = \mathscr{S}\left(\mathbb{A}^{(n)}\right)$ is a totally symmetric tensor with components

$$\widetilde{A}^{(n)}_{\alpha_1\ldots\alpha_n} = \frac{1}{n!}\sum_{\pi(\alpha_1\ldots\alpha_n)} A^{(n)}_{\alpha_1\ldots\alpha_n}.$$

As we will see shortly, there also exists an operator that makes a totally symmetric tensor totally symmetric and traceless.

5.2.3 Maxwell-Cartesian spherical harmonics

It is well known that any solid spherical harmonic of degree n, when expressed in Cartesian coordinates: $S_n(\mathbf{v}) = S_n(v_x, v_y, v_z) = v^n Y_n(\mathbf{v}/v)$, is a harmonic polynomial, homogeneous of degree n ([49, Thm. 3.67]), i.e.

$$\nabla^2 S_n(\mathbf{v}) = 0, \quad S_n(\lambda\mathbf{v}) = \lambda^n S_n(\mathbf{v})$$

(in fact, there is an alternative definition of solid spherical harmonic as a function satisfying these two properties and of surface spherical harmonic as a restriction

of such function to the unit sphere). The space of such polynomials has dimension $2n+1$ [2] and is isomorphic to the space of TST tensors of rank n ([49, Chap. 3]). In particular, when the components of such a tensor do not depend on v_x, v_y, v_z, (or Ω_x, Ω_y, Ω_z) they comprise coefficients of the solid (or surface) spherical harmonic and therefore also (in the $\mathcal{S}_2$-restricted case) the coefficients $\phi_n^m(\mathbf{r}, E)$ in the expansion (5.2) (when transformed back to spherical coordinates). Hence, by solving the set of tensorial equations with these tensors of ranks $n \leq N$ as unknowns, we accomplish essentially the same thing as by solving the ordinary P$_N$ equations, but we may additionally utilize the TST tensor character of the equations.

To construct this set, we will build on the ideas presented by Johnston ([64]) and Applequist ([6]). Johnston arrived at a "far more symmetric and compact" ([64, p. 1455]) form of the angularly discretized general Boltzmann-Vlasov equation by expanding the flow function in terms of *Maxwell-Cartesian spherical harmonic tensors* rather than the usual surface spherical harmonics. The term "Maxwell-Cartesian spherical harmonic tensor" has been coined by Applequist in [7], who in his older paper [6] showed a systematic way of obtaining a special TST tensor of any rank n whose components are spherical harmonics of degree n in Cartesian frame of reference as defined by Maxwell in [81, p. 160]. Specifically, in present notation, he showed that Maxwell's spherical harmonics based on Cartesian axes can be obtained (up to a normalization constant) as components of

$$\mathscr{D}\mathbf{v}^n \quad (\text{or } \mathscr{D}\boldsymbol{\Omega}^n)$$

where $\mathscr{D}$ is the so-called *detracer operator* which projects a general totally symmetric tensor of rank n into the space of totally symmetric and traceless tensors of rank n. The result of applying the detracer operator on a totally symmetric

[2] Homogeneous polynomial of degree n has $\frac{(n+1)(n+2)}{2}$ coefficients; since $\nabla^2 S_n(\mathbf{v})$ is a homogeneous polynomial of degree $n-2$ with $\frac{n(n-1)}{2}$ coefficients expressed in terms of those of S_n, the condition $\nabla^2 S_n(\mathbf{v}) = 0$ reduces the number of independent coefficients of S_n to $\frac{(n+1)(n+2)}{2} - \frac{n(n-1)}{2} = 2n+1$.

tensor $\mathbb{A}^{(n)}$ is a TST tensor with components:

$$\mathscr{D}A^{(n)}_{\alpha_1\dots\alpha_n} = \sum_{m=0}^{\lfloor n/2\rfloor} \frac{(-1)^m(2n-2m-1)!!}{(2n-1)!!} \sum_{\{\pi(\alpha_1\dots\alpha_n)\}_{\mathrm{d}}} \delta_{\alpha_1\alpha_2}\cdots\delta_{\alpha_{2m-1}\alpha_{2m}} A^{(n)}_{\beta_1\beta_1\dots\beta_m\beta_m\alpha_{2m+1}\dots\alpha_n}, \tag{5.7}$$

where $\{\pi(\alpha_1\dots\alpha_n)\}_{\mathrm{d}}$ denotes the set of all permutations that give distinct terms in the last sum (considering their total symmetry) and

$$(2n-1)!! = (2n-1)(2n-3)\cdots 3\cdot 1 \quad \text{with} \quad (-1)!! = 1.$$

The proof that $\mathscr{D}A^{(n)}_{\alpha_1\dots\alpha_n}$ is a traceless tensor is given in [6, Sec. 5] and we will not repeat it here as it is straightforward but technically involved. We note however that we include the factor $(2n-1)!!$ in the denominator in (5.7) to make $\mathscr{D}$ idempotent (and thus a true projection).

Being general surface spherical harmonics according to Def. 3, Maxwell's surface spherical harmonics are linear combination of spherical harmonics and hence share many of the properties of the latter. An extensive presentation of these properties is given in [6, 7] and we will recall some of them in further sections. We will use the following form of Maxwell-Cartesian surface spherical harmonics of degree n as components of rank-n TST tensors (called sometimes *Maxwell-Cartesian tensors*):

$$\mathbb{P}^{(n)}(\mathbf{\Omega}) = \mathscr{D}\mathbf{\Omega}^n \tag{5.8}$$

which differ from those used by Applequist in that they have coefficient 1 at their first term but which coincide with those used by Johnston.

We note that projection of $\mathbb{P}^{(n)}$ along, say, z-axis (or any other because of the symmetry) yields (up to a normalization factor) the Legendre polynomials:

$$\mathbb{P}^{(n)}(\mathbf{\Omega})\cdot\mathbf{e}_z^n = P^{(n)}_{\alpha_1\dots\alpha_n}\delta_{3\alpha_1}\cdots\delta_{3\alpha_n} = P^{(n)}_{33\dots3} = \frac{n!}{(2n-1)!!}P_n(\mu) \equiv C_nP_n(\mu) \tag{5.9}$$

and hence components of $\mathbb{P}^{(n)}$ could be explicitly obtained by extending the well-known formulas for Legendre polynomials, for instance

$$\begin{aligned}\mathbb{P}^{(n)}(\mathbf{\Omega}) = \mathbf{\Omega}^n &- \frac{n(n-2)}{2(2n-1)}\mathscr{S}\left(\mathbb{I}\otimes\mathbf{\Omega}^{n-2}\right)\\ &+ \frac{n(n-1)(n-2)(n-3)}{8(2n-1)(2n-3)}\mathscr{S}\left(\mathbb{I}^2\otimes\mathbf{\Omega}^{n-4}\right)\mp\dots\end{aligned} \tag{5.10}$$

([17, Chap. VI], [64]) First few Maxwell-Cartesian surface harmonics generated by this formula are shown in the table below together with the corresponding spherical harmonics and Legendre polynomials.

n	$\mathbf{Y}_n(\mathbf{\Omega}) \propto$	$\mathbb{P}^{(n)}(\mathbf{\Omega})$	$C_n P_n(\mu)$
0	1	1	1
1	Ω_x, Ω_y, Ω_z	Ω_x, Ω_y, Ω_z	μ
2	$-\Omega_x^2 - \Omega_y^2 + 2\Omega_z^2$, $\Omega_y\Omega_z$, $\Omega_z\Omega_x$, $\Omega_x\Omega_y$, $\Omega_x^2 - \Omega_y^2$	$\Omega_x^2 - \frac{1}{3}$, $\Omega_y^2 - \frac{1}{3}$, $\Omega_z^2 - \frac{1}{3}$, $\Omega_x\Omega_y$, $\Omega_x\Omega_z$, $\Omega_y\Omega_z$	$\mu^2 - \frac{1}{3}$

Table 5.1: Spherical harmonics, Maxwell-Cartesian surface harmonics and Legendre polynomials up to degree $n = 2$.

The symmetric nature of the Maxwell-Cartesian surface harmonics as opposed to the ordinary spherical harmonics can be seen from the third row in the table. We also note that since there are $\frac{(n+1)(n+2)}{2}$ distinct components in a general totally symmetric tensor of rank n, the number of Maxwell-Cartesian harmonics is greater than the number of spherical harmonics of the same degree, which is $(2n + 1)$ for $n \geq 2$. However, due to the $\frac{n(n-1)}{2}$ conditions arising from the tracelessness property, there are just $2n + 1$ independent components in $\mathbb{P}^{(n)}$ (see also the footnote on pg. 90). This also indicates the fact that unlike the spherical harmonics with same degree but different orders, components of $\mathbb{P}^{(n)}$ are not all $L^2(\mathcal{S}_2)$-orthogonal. Nevertheless, Maxwell-Cartesian surface tensors of different degrees are orthogonal in the following sense:

$$\int_{\mathcal{S}_2} \mathbb{P}^{(n)}(\mathbf{\Omega}) \otimes \mathbb{P}^{(m)}(\mathbf{\Omega})\, \mathrm{d}\mathbf{\Omega} = \mathbb{O}^{m+n}, \quad n \neq m. \tag{5.11}$$

The Maxwell-Cartesian tensors have been actively used for solving various electro-magnetics and quantum-mechanical problems (see the references in [7]), but there is apparently only one attempt to bring them to the field of neutron transport. In a recent article [27], Coppa re-derived the Maxwell-Cartesian surface spherical harmonic tensors from their analogy with Legendre polynomials (see above) and the requirement that they satisfy the addition theorem of the form (5.3). The paper then proceeds by multiplying the NTE with these tensors

and integrating over the sphere to obtain a system of tensorial equations, each of which being itself a symmetric set of PDE's determining the components of the flux moment tensor defined by

$$\psi^{(n)}(\mathbf{r}) = \int_{\mathcal{S}_2} \psi(\mathbf{r}, \mathbf{\Omega}) \mathbb{P}^{(n)}(\mathbf{\Omega}) \, \mathrm{d}\mathbf{\Omega} \,. \qquad ([27, \text{Eqn. } (41)])$$

Note that this definition implies that the flux moment tensors $\psi^{(n)}$ are totally symmetric and traceless. However, only the symmetry property appears to be recognized in [27]. Using rather complicated tensor relations, the system is constructed again by analogy with the procedure leading to the P_N system. This could be misleading, however, as there is no indication of in what sense the moments obtained by solving the set could be interpreted as coefficients in an expansion of angular flux. Note that in the usual P_N method where the angular flux is expanded into a series of spherical harmonics, the key ingredient that allows us to deduce the system of equations analogous to [27, Eqn. (48)]) for the coefficients in the expansion is orthogonality (and hence linear independence) of all Y_n^m. We do not have this property in the set of Maxwell-Cartesian harmonics (which does not seem to be taken into account in [27]).

This missing part of Coppa's derivation can be filled-in by taking into account the tracelessness property of the Maxwell-Cartesian tensors. This will be demonstrated in the following section using the results of Johnston ([64]) and Applequist ([6]). We will see that this more fundamental approach leads to a set of equations equivalent with [27, Eqn. (42)]. We will also see a clear connection to the 1D P_N equations (based on the expansion of angular flux in Legendre polynomials) which could be expected from the formal equivalence of the Maxwell-Cartesian tensors and Legendre polynomials mentioned above ((5.10), Tab. 5.1). Motivated by this observation, we will finally exhibit the connection between the derived equations and the SP_3 equations in Sec. 5.4.1[3].

[3] Analogous relationship with the P_1 and SP_2 equations have been shown in [27].

5.3 Derivation of the MCP$_N$ approximation

5.3.1 First attempts

Using the results of Sec. 3.3.3, the expansion of angular neutron flux in terms of surface spherical harmonics can be written as[4]

$$\psi(\mathbf{r},\mathbf{\Omega}) = \sum_{n=0}^{\infty} \frac{2n+1}{4\pi} \int_{\mathcal{S}_2} \psi(\mathbf{r},\mathbf{\Omega}') P_n(\mathbf{\Omega}\cdot\mathbf{\Omega}')\,\mathrm{d}\mathbf{\Omega}\,. \tag{5.12}$$

As shown in [6, Sec. 7, Corollary II], Maxwell-Cartesian tensors satisfy the following form of addition theorem:

$$\mathbb{P}^{(n)}(\mathbf{\Omega})\cdot\mathbb{P}^{(n)}(\mathbf{\Omega}') = C_n P_n(\mathbf{\Omega}\cdot\mathbf{\Omega}') \tag{5.13}$$

(C_n defined in (5.9)). Combining the two results, we obtain

$$\psi(\mathbf{r},\mathbf{\Omega}) = \sum_{n=0}^{\infty} \frac{2n+1}{4\pi C_n} \left[\int_{\mathcal{S}_2} \psi(\mathbf{r},\mathbf{\Omega}')\mathbb{P}^{(n)}(\mathbf{\Omega}')\,\mathrm{d}\mathbf{\Omega}'\right]\cdot\mathbb{P}^{(n)}(\mathbf{\Omega}). \tag{5.14}$$

If we now define the *n-th angular flux moment tensor* by

$$\psi^{(n)}(\mathbf{r}) \coloneqq \frac{2n+1}{4\pi C_n} \int_{\mathcal{S}_2} \psi(\mathbf{r},\mathbf{\Omega})\mathbb{P}^{(n)}(\mathbf{\Omega})\,\mathrm{d}\mathbf{\Omega} \tag{5.15}$$

equation (5.14) becomes

$$\psi(\mathbf{r},\mathbf{\Omega}) = \sum_{n=0}^{\infty} \psi^{(n)}(\mathbf{r})\cdot\mathbb{P}^{(n)}(\mathbf{\Omega}). \tag{5.16}$$

Note that we have, similarly to the P$_N$ case,

$$\phi = 4\pi\psi^{(0)}, \quad \mathrm{J} = \frac{4\pi}{3}\begin{bmatrix}\psi_x^{(1)}\\ \psi_y^{(1)}\\ \psi_z^{(1)}\end{bmatrix}.$$

[4] with the equality considered in the sense that

$$f = \sum_{n=0}^{\infty} g_n \;\Leftrightarrow\; \lim_{n\to\infty}\left\| f - \sum_{k=0}^{n} g_k \right\|_{L^2(\mathcal{S}_2)} = 0$$

Although equation (5.16) with (5.15) looks like a (generalized) Fourier series, we should keep in mind that

$$\int_{\mathcal{S}_2} \mathbb{P}^{(m)}(\mathbf{\Omega}) \otimes \sum_{n=0}^{\infty} \psi^{(n)}(\mathbf{r}) \cdot \mathbb{P}^{(n)}(\mathbf{\Omega}) \, \mathrm{d}\mathbf{\Omega} \neq \psi^{(m)}$$

in general and regard (5.15) as definition.

We analogously expand also the volumetric source term, leading to

$$q(\mathbf{r}, \mathbf{\Omega}) = \sum_{n=0}^{\infty} q^{(n)}(\mathbf{r}) \cdot \mathbb{P}^{(n)}(\mathbf{\Omega}) \tag{5.17}$$

with

$$q^{(n)}(\mathbf{r}) \coloneqq \frac{2n+1}{4\pi C_n} \int_{\mathcal{S}_2} q(\mathbf{r}, \mathbf{\Omega}) \mathbb{P}^{(n)}(\mathbf{\Omega}) \, \mathrm{d}\mathbf{\Omega} \, .$$

To find the relations that must be satisfied by the angular expansion moments $\psi^{(n)}(\mathbf{r})$ in order for (5.16) (or equivalently (5.12)) to be the solution of the transport equation (5.1), we insert the expansion (5.16) into (5.1) (with source term represented according to (5.17)). Let us first look at the transfer part. Applying eq. (5.13) in the expansion (5.4), we immediately simplify the scattering term[5]:

$$\begin{aligned}
\int_{\mathcal{S}_2} \sigma_s(\mathbf{r}, \mathbf{\Omega} \cdot \mathbf{\Omega}') \psi(\mathbf{r}, \mathbf{\Omega}') \, \mathrm{d}\mathbf{\Omega}' &= \sum_{n=0}^{\infty} \frac{2n+1}{4\pi} \sigma_{sn}(\mathbf{r}) \int_{\mathcal{S}_2} P_n(\mathbf{\Omega}' \cdot \mathbf{\Omega}) \psi(\mathbf{r}, \mathbf{\Omega}') \, \mathrm{d}\mathbf{\Omega}' \\
&= \sum_{n=0}^{\infty} \frac{2n+1}{4\pi C_n} \sigma_{sn}(\mathbf{r}) \left[\int_{\mathcal{S}_2} \mathbb{P}^{(n)}(\mathbf{\Omega}') \psi(\mathbf{r}, \mathbf{\Omega}') \, \mathrm{d}\mathbf{\Omega}' \right] \cdot \mathbb{P}^{(n)}(\mathbf{\Omega}) \\
&= \sum_{n=0}^{\infty} \sigma_{sn}(\mathbf{r}) \psi^{(n)}(\mathbf{r}) \cdot \mathbb{P}^{(n)}(\mathbf{\Omega}).
\end{aligned}$$

Using def. (5.15), the fission part will has the expected form:

$$\int_{\mathcal{S}_2} \frac{\nu\sigma_f(\mathbf{r})}{4\pi} \psi(\mathbf{r}, \mathbf{\Omega}') \, \mathrm{d}\mathbf{\Omega}' = \frac{\nu\sigma_f(\mathbf{r})}{4\pi} \int_{\mathcal{S}_2} \psi(\mathbf{r}, \mathbf{\Omega}') \, \mathrm{d}\mathbf{\Omega}' = \nu\sigma_f(\mathbf{r}) \phi(\mathbf{r}). \tag{5.18}$$

Therefore, by inserting (5.16) into (5.1) and using these results we obtain

$$\sum_{n=0}^{\infty} \left[\mathbf{\Omega} \cdot \nabla \psi^{(n)} + \sigma_t \psi^{(n)} - \sigma_{sn} \psi^{(n)} - \delta_{n0} \nu\sigma_f \phi - q^{(n)} \right] \cdot \mathbb{P}^{(n)}(\mathbf{\Omega}) = 0 \tag{5.19}$$

[5]In practice, the scattering cross-section moments σ_{sn} are available only up to a certain anisotropy degree $N_s \leq \infty$; in that case we set $\sigma_{sn} = \sigma_{sk}$ for $n \leq N_s$ and $\sigma_{sn} = 0$ otherwise.

where each term in the brackets is dependent only on $\mathbf{r}$ (which is omitted for brevity).

Formulation of the moment equations for $\psi^{(n)}(\mathbf{r})$ is now hampered by linear dependence among certain functions in each $\mathbb{P}^{(n)}(\mathbf{\Omega})$ (so that we cannot deduce from (5.19) that for each n, all components of the expression in brackets must vanish), as well as by the advection term which still contains $\mathbf{\Omega}$. Both issues may be however overcome by the so-called *detracer exchange theorem* ([7, Sec. 5.2]).

Theorem 7. *If $\mathbb{A}^{(n)}$ and $\mathbb{B}^{(n)}$ are totally symmetric tensors of rank n, then*

$$\mathbb{A}^{(n)} \cdot \mathscr{D}\mathbb{B}^{(n)} = \mathbb{B}^{(n)} \cdot \mathscr{D}\mathbb{A}^{(n)}.$$

Proof. The theorem easily follows from the definitions of the detracer operator (5.7) and tensor contraction. □

Since $\psi^{(n)}(\mathbf{r})$ is by definition totally symmetric and traceless and $\mathbf{\Omega}^n$ totally symmetric, using Theorem 7 with $\mathbb{A}^{(n)} \equiv \psi^{(n)}(\mathbf{r})$ and $\mathbb{B}^{(n)} \equiv \mathbf{\Omega}^n$ and the definition (5.8) of Maxwell-Cartesian tensors shows that[6]

$$\psi^{(n)}(\mathbf{r}) \cdot \mathbb{P}^{(n)}(\mathbf{\Omega}) = \psi^{(n)}(\mathbf{r}) \cdot \mathbf{\Omega}^n$$

and the expansion (5.16) is thus equivalent to a power series in $\mathbf{\Omega}$:

$$\psi(\mathbf{r}, \mathbf{\Omega}) = \sum_{n=0}^{\infty} \psi^{(n)}(\mathbf{r}) \cdot \mathbf{\Omega}^n$$

(similarly for (5.17)). The advection term therefore simplifies as follows:

$$(\mathbf{\Omega} \cdot \nabla)\psi^{(n)} \cdot \mathbf{\Omega}^n = \begin{Bmatrix} \Omega_\beta \nabla_\beta \psi^{(n)}_{\alpha_0,\ldots,\alpha_{n-1}} \Omega_{\alpha_{n-1}} \cdots \Omega_{\alpha_0} & \text{if } n \geq 1, \\ \Omega_\beta \nabla_\beta \phi & \text{if } n = 0 \end{Bmatrix} = \nabla\psi^{(n)} \cdot \mathbf{\Omega}^{n+1}$$

for $n = 0, 1, \ldots$ and by reindexing:

$$(\mathbf{\Omega} \cdot \nabla)\psi^{(n-1)} \cdot \mathbf{\Omega}^{n-1} = \nabla\psi^{(n-1)} \cdot \mathbf{\Omega}^n$$

for $n = 1, 2, \ldots$. We can therefore rewrite eq. (5.19) as

$$\sum_{n=0}^{\infty} \left[\nabla\psi^{(n-1)} + \sigma_t \psi^{(n)} - \sigma_{sn}\psi^{(n)} - \delta_{n0}\nu\sigma_f\phi - q^{(n)}\right] \cdot \mathbf{\Omega}^n = 0 \tag{5.20}$$

[6]Recall that the detracer is a projection operator into the space of TST tensors and thus leaves the already TST tensor unchanged.

with the term with negative rank discarded. Although eq. (5.20) closely resembles a vanishing linear combination of monomials [7]

$$\mathbf{v}^n = x^{n_1} y^{n_2} z^{n_3}, \quad n_1 + n_2 + n_3 = n$$

whose linear independence would imply that the expression in brackets must be a zero tensor for each n, by examining the not fully sensible equations that we would obtain in this way we might suspect that this is not the case. The reason is given in the following section.

5.3.2 Linear independence of monomials restricted to $\mathcal{S}_2$

When monomials $\mathbf{v}^n$ of different degrees are restricted to the unit sphere, there appear nontrivial but vanishing linear combinations of them (consider e.g. $\Omega_x^2 + \Omega_y^2 + \Omega_z^2 - 1$). However, when only the components of $\mathbf{\Omega}^n$ for a *single* n are considered, they *are* linearly independent according to the following lemma.

Lemma 5. *Let $n \in \mathbb{N}_0$ be fixed. If $\mathbb{A}^{(n)}$ is a totally symmetric tensor of rank n, independent of $\mathbf{\Omega}$, and*

$$\mathbb{A}^{(n)} \cdot \mathbf{\Omega}^n = 0, \tag{5.21}$$

then $\mathbb{A}^{(n)} = \mathbb{O}^{(n)}$.

Proof. If $\mathbb{A}^{(n)}$ is a totally symmetric tensor, then

$$\mathbb{A}^{(n)} \cdot \mathbf{\Omega}^n = \frac{1}{v^n} \mathbb{A}^{(n)} \cdot \mathbf{v}^n$$

represents a linear combination of monomials of degree n, with coefficients being the components of $\mathbb{A}^{(n)}$. By linear independence of these monomials we conclude that if the equality (5.21) holds, then all components of $\mathbb{A}^{(n)}$ must vanish. □

One possibility how to extend the above result to an arbitrary linear combination of direction tensors with different ranks is to require the moment tensors to be not only totally symmetric, but also traceless.

[7] Since we will not need to distinguish between the position vector $\mathbf{r}$ and the velocity vector $\mathbf{v}$ here, we use the standard Cartesian coordinates $[x, y, z]$ for the latter as well.

Theorem 8. *If $\mathbb{A}^{(n)}$ is a TST tensor independent of $\mathbf{\Omega}$ for all $n = 0, 1, \ldots$ and*

$$\sum_{n=0}^{\infty} \mathbb{A}^{(n)} \cdot \mathbf{\Omega}^n = 0,$$

then $\mathbb{A}^{(n)} = \mathbb{O}^{(n)}$ for all n.

Proof. As discussed in Sec. 5.2.3, any TST tensor $\mathbb{A}^{(n)}$ uniquely identifies a surface spherical harmonic of degree n by

$$\mathcal{Y}_n(\mathbf{\Omega}) = \mathbb{A}^{(n)} \cdot \mathbf{\Omega}^n.$$

Since surface spherical harmonics of different degrees are linearly independent,

$$\sum_{n=0}^{\infty} \mathcal{Y}_n(\mathbf{\Omega}) = 0 \quad \Rightarrow \quad \mathcal{Y}_n(\mathbf{\Omega}) = 0 \ \ \forall n.$$

The conclusion follows from the above lemma. □

5.3.3 The MCP$_N$ approximation

Using Thm. 8, we see that we can obtain from (5.20) a set of moment equations if we ensure that the expression in square brackets is a TST tensor for each n. Since $\psi^{(n)}$ and $q^{(n)}$ are TST by definition, we only have to symmetrize and detrace the advection terms. We need to be careful, however, not to change the original equation. This can be done by a clever rearranging of the terms in the sum.

First note that thanks to the symmetry of $\mathbf{\Omega}^n$ we obviously have

$$\nabla\psi^{(n-1)} \cdot \mathbf{\Omega}^n = \mathscr{S}\left(\nabla\psi^{(n-1)}\right) \cdot \mathbf{\Omega}^n.$$

The symmetrized advection terms can be put into an equivalent TST form by

$$\mathscr{S}\left(\nabla\psi^{(n-1)}\right) = \mathscr{D}\mathscr{S}\left(\nabla\psi^{(n-1)}\right) - \left(\mathscr{D}\mathscr{S}\left(\nabla\psi^{(n-1)}\right) - \mathscr{S}\left(\nabla\psi^{(n-1)}\right)\right).$$

The result of the detracer operation on $\mathscr{S}\left(\nabla\psi^{(n-1)}\right)$ can be seen by using

$$A^{(n)}_{\alpha_1\ldots\alpha_n} = \mathscr{S}\left(\nabla_{\alpha_1}\psi^{(n-1)}_{\alpha_2\ldots\alpha_n}\right), \quad n \geq 2$$

in eq. (5.7). Taking into account the tracelessness of $\psi^{(n-1)}$, the only nonzero terms in the outer sum will be those with $m = 0$ and $m = 1$. For $m = 0$, we

obtain just the original tensor $\mathbb{A}^{(n)} = \mathscr{S}\left(\nabla\psi^{(n-1)}\right)$, while for $m = 1$, we obtain the inner sum over terms

$$\mathscr{S}\left(\delta_{\alpha_1\alpha_2}\nabla_\beta\psi^{(n-1)}_{\beta\alpha_3\ldots\alpha_n}\right), \quad \mathscr{S}\left(\delta_{\alpha_1\alpha_3}\nabla_\beta\psi^{(n-1)}_{\alpha_2\beta\alpha_4\ldots\alpha_n}\right), \quad \text{etc.},$$

that is $n-1$ times $\mathscr{S}\left(\mathbb{I}\otimes\nabla\cdot\psi^{(n-1)}\right)$. We therefore get

$$\begin{aligned}\nabla\psi^{(n-1)}\cdot\boldsymbol{\Omega}^n &= \mathscr{S}\left(\nabla\psi^{(n-1)}\right)\cdot\boldsymbol{\Omega}^n \\ &= \mathscr{S}\left(\nabla\psi^{(n-1)} - \frac{n-1}{2n-1}\mathbb{I}\otimes\nabla\cdot\psi^{(n-1)} + \frac{n-1}{2n-1}\mathbb{I}\otimes\nabla\cdot\psi^{(n-1)}\right)\cdot\boldsymbol{\Omega}^n \\ &= \mathscr{S}\left(\nabla\psi^{(n-1)} - \frac{n-1}{2n-1}\mathbb{I}\otimes\nabla\cdot\psi^{(n-1)}\right)\cdot\boldsymbol{\Omega}^n \\ &+ \quad \frac{n-1}{2n-1}\left(\nabla\cdot\psi^{(n-1)}\right)\cdot\boldsymbol{\Omega}^{n-2}\end{aligned} \tag{5.22}$$

where the last equality comes from the realization that $\mathbb{I}\cdot\boldsymbol{\Omega}^2 = \boldsymbol{\Omega}\cdot\boldsymbol{\Omega} = 1$ and $\nabla\cdot\psi^{(n-1)}$ is totally symmetric since $\psi^{(n-1)}$ is. Note that we have actually obtained that

$$\nabla\psi^{(n-1)}\cdot\boldsymbol{\Omega}^n = \mathscr{D}\mathscr{S}\left(\nabla\psi^{(n-1)}\right)\cdot\boldsymbol{\Omega}^n + \frac{n-1}{2n-1}\left(\nabla\cdot\psi^{(n-1)}\right)\cdot\boldsymbol{\Omega}^{n-2}. \tag{5.23}$$

Eq. (5.23) shows that even though $\nabla\psi^{(n-1)}$ by itself is neither symmetric nor traceless, its contraction with $\boldsymbol{\Omega}^n$ can be split into a sum of two contractions of TST tensors (total symmetry and tracelessness of the latter follows from the fact that $\psi^{(n-1)}$ is TST). By introducing this splitting into eq. (5.20) and regrouping the sum by $\boldsymbol{\Omega}^n$, we finally obtain

$$\begin{aligned}\sum_{n=0}^{\infty}\Bigg\{\mathscr{S}\left(\nabla\psi^{(n-1)} - \frac{n-1}{2n-1}\mathbb{I}\otimes\nabla\cdot\psi^{(n-1)}\right) + \frac{n+1}{2n+3}\nabla\cdot\psi^{(n+1)} + \\ + \sigma_t\psi^{(n)} - \sigma_{sn}\psi^{(n)} - \delta_{n0}\nu\sigma_f\phi - q^{(n)}\Bigg\}\cdot\boldsymbol{\Omega}^n = 0\end{aligned} \tag{5.24}$$

(again with the non-sensical tensors with negative ranks discarded).

Equation (5.24) is completely equivalent to eq. (5.20) or (5.19), but regrouped to a form of vanishing linear combination of $\boldsymbol{\Omega}^n$ with TST coefficient tensors for each n. Theorem 8 therefore applies and we get the desired set of first-order

partial differential equations for the angular flux moments:

$$\begin{aligned}\frac{n+1}{2n+3}\nabla\cdot\psi^{(n+1)}+\mathscr{S}\left(\nabla\psi^{(n-1)}-\frac{n-1}{2n-1}\mathbb{I}\otimes\nabla\cdot\psi^{(n-1)}\right)\\ +\sigma_t\psi^{(n)}-\sigma_{sn}\psi^{(n)}-\delta_{n0}\nu\sigma_f\phi=q^{(n)},\qquad n=0,1,\dots\end{aligned}\tag{5.25}$$

Note that in passing to a finite approximation by the simplest closure $\nabla\cdot\psi^{(N+1)}\equiv\mathbb{O}^{(N)}$ for some $N\geq 0$, we do not spoil the TST character of the N-th coefficient tensor (in view of (5.22), this closure actually means that we are neglecting the nonzero traces of $\nabla\psi^{(N-1)}$). The set (5.25) for $n\leq N$ may be regarded as an alternative to the ordinary P$_N$ equations and because its solution represents the expansion of angular flux into Maxwell-Cartesian surface spherical harmonics of degrees up to N, it will be called *MCP$_N$ approximation*.

5.4 The MCP$_3$ equations

We now explicitly state the first four tensor equations in the MCP$_N$ set. Let

$$\Sigma_n=\sigma_t-\sigma_{sn}-\delta_{n0}\nu\sigma_f.$$

Then the MCP$_3$ equations read

$$\begin{aligned}\tfrac{1}{3}\nabla\cdot\psi^{(1)}+\Sigma_0\phi&=q^{(0)}\\ \tfrac{2}{5}\nabla\cdot\psi^{(2)}+\nabla\phi+\Sigma_1\psi^{(1)}&=q^{(1)}\\ \tfrac{3}{7}\nabla\cdot\psi^{(3)}+\mathscr{S}\left(\nabla\otimes\psi^{(1)}\right)-\tfrac{1}{3}\mathbb{I}\otimes\nabla\cdot\psi^{(1)}+\Sigma_2\psi^{(2)}&=q^{(2)}\\ \mathscr{S}\left(\nabla\otimes\psi^{(2)}\right)-\tfrac{2}{5}\mathscr{S}\left(\mathbb{I}\otimes\nabla\cdot\psi^{(2)}\right)+\Sigma_3\psi^{(3)}&=q^{(3)}\end{aligned}\tag{5.26}$$

(where we included the $\otimes$ symbol whenever multiplication of tensors of rank ≥ 1 occurred). We can see that the MCP$_1$ set is equivalent to the P$_1$ set. For $N\geq 3$, we have to take into account the definition of the moments as *traceless* totally symmetric tensors. If we considered general symmetric moment tensors

$$\psi^{(2)}=\begin{bmatrix}\psi^{(2)}_{11}&\psi^{(2)}_{12}&\psi^{(2)}_{13}\\ \psi^{(2)}_{12}&\psi^{(2)}_{22}&\psi^{(2)}_{23}\\ \psi^{(2)}_{13}&\psi^{(2)}_{23}&\psi^{(2)}_{33}\end{bmatrix}\tag{5.27}$$

and similarly for $\psi^{(3)}$, etc., we would obtain more linearly independent equations than in the corresponding P$_N$ set. However, by restricting to TST tensors (e.g. by using the detraced versions of (5.27) and higher moments:

$$\mathscr{D}\psi^{(2)} = \begin{bmatrix} \frac{1}{3}(2\psi_{11}^{(2)} - \psi_{22}^{(2)} - \psi_{33}^{(2)}) & \psi_{12}^{(2)} & \psi_{13}^{(2)} \\ \psi_{12}^{(2)} & \frac{1}{3}(-\psi_{11}^{(2)} + 2\psi_{22}^{(2)} - \psi_{33}^{(2)}) & \psi_{23}^{(2)} \\ \psi_{13}^{(2)} & \psi_{23}^{(2)} & \frac{1}{3}(-\psi_{11}^{(2)} - \psi_{22}^{(2)} + 2\psi_{33}^{(2)}) \end{bmatrix})$$

we obtain the same number of linearly independent equations in the set (5.24) as in the P$_N$ set. Moreover, these equations can be written in the form of (3.10):

$$\mathbf{A}^x_{\mathrm{MCP}_3}\frac{\partial\Phi}{\partial x} + \mathbf{A}^y_{\mathrm{MCP}_3}\frac{\partial\Phi}{\partial y} + \mathbf{A}^z_{\mathrm{MCP}_3}\frac{\partial\Phi}{\partial z} + \left[\sigma_t\mathbf{I} - \mathbf{K}_{\mathrm{MCP}_3}\right]\Phi = \mathrm{Q}, \tag{5.28}$$

where Φ and Q now contain the different MCP$_3$ moments – e.g. Φ is a column vector with components

$$\phi,$$

$$\psi_1^{(1)}, \psi_2^{(1)}, \psi_3^{(1)},$$

$$\psi_{11}^{(2)}, \psi_{12}^{(2)}, \psi_{13}^{(2)}, \psi_{22}^{(2)}, \psi_{23}^{(2)}, \psi_{33}^{(2)},$$

$$\psi_{111}^{(3)}, \psi_{211}^{(3)}, \psi_{311}^{(3)}, \psi_{221}^{(3)}, \psi_{321}^{(3)}, \psi_{331}^{(3)}, \psi_{222}^{(3)}, \psi_{322}^{(3)}, \psi_{332}^{(3)}, \psi_{333}^{(3)}.$$

The advection matrices $\mathbf{A}^x_{\mathrm{MCP}_3}$, $\mathbf{A}^y_{\mathrm{MCP}_3}$, $\mathbf{A}^z_{\mathrm{MCP}_3}$ have the same eigenvalues, which are moreover equal to those of the advection matrices of the P$_N$ approximation (only the multiplicity of the zero eigenvalue is greater, reflecting the linear dependence between the MCP$_3$ equations) – see App. D. The paper [6] shows how to systematically select linearly independent subsets of the TST tensors, which could be used to actually solve the MCP$_3$ equations. However, we stress that this is not our intention since this would be no easier than solving the ordinary P$_N$ equations. Rather, we will try to exploit the symmetric and traceless structure of the MCP$_N$ equations to provide a new perspective on the SP$_N$ approximation presented in Chap. 4.

When each moment tensor is projected along a chosen axis (say z), we obtain one dimensional equations which are exactly the same as the 1D P$_N$ equations (since both symmetrized terms in every $n \geq 2$ MCP$_N$ equation become equal

to $\frac{d\psi_z^{(n-1)}}{dz}$) if we take into account the definition of moments in the classical P$_N$ equations and multiply each $\psi_z^{(n)}$ accordingly by a factor $2n+1$ (similarly, by multiplying each $\psi^{(n)}$ by $\frac{n!}{(2n-1)!!}$, we arrive at the system derived by Coppa [27]). This indicates the possibility to investigate the original ad-hoc derivation of the SP$_N$ equation by formal extension of the 1D P$_N$ equations into 3D in the current tensorial framework.

5.4.1 Reduction of the MCP$_3$ system

In an analogy to the simplified P$_N$ approach, we start this investigation by reducing the system (5.26) to two equations governing the even-order tensorial moments $\psi^{(0)}$ and $\psi^{(2)}$. We need to assume sufficient differentiability of the angular flux and source moments in eq. (5.26) and no void regions (i.e., all $\Sigma_n > 0$). For simplicity, let us further focus on the case of isotropic volumetric sources and scattering:

$$q^{(1)} \equiv q, \quad \Sigma_0 = \sigma_t - \sigma_{s0} - \nu\sigma_f = \sigma_a - \nu\sigma_f,$$
$$q^{(n)} = 0, \quad \Sigma_n = \sigma_t \equiv \Sigma, \qquad n = 1, 2, \ldots$$

(generalization to higher anisotropy degrees is straightforward but technically more involved).

As a starting point, let us take the system (5.26) in which the above assumptions are taken into account and the substitution

$$\psi^{(n)} \longrightarrow \frac{n!}{(2n-1)!!}\psi^{(n)}$$

is made for convenience:

$$\nabla \cdot \psi^{(1)} + \Sigma_0 \phi = q \tag{5.29}$$

$$-\nabla \cdot \psi^{(2)} - \tfrac{1}{3}\nabla\phi = \Sigma\psi^{(1)} \tag{5.30}$$

$$-\nabla \cdot \psi^{(3)} - \tfrac{2}{5}\mathscr{S}\left(\nabla \otimes \psi^{(1)}\right) + \tfrac{2}{15}\mathbb{I} \otimes \nabla \cdot \psi^{(1)} = \Sigma\psi^{(2)} \tag{5.31}$$

$$-\tfrac{3}{7}\mathscr{S}\left(\nabla \otimes \psi^{(2)}\right) + \tfrac{6}{35}\mathscr{S}\left(\mathbb{I} \otimes \nabla \cdot \psi^{(2)}\right) = \Sigma\psi^{(3)} \tag{5.32}$$

To further simplify the use of basic tensor identities from App. E, we consider the equations in a homogeneous region with constant Σ and Σ_0 [8]. The deriva-

[8]Note that this does not preclude heterogeneous subregions with the physical cross-sections σ_a, σ_{s0} and $\nu\sigma_f$ varying in such a way that $\sigma_a - \nu\sigma_f$ and $\sigma_a + \sigma_s$ are constant.

tion below obviously extends to the case of sufficiently smooth cross-sections. In practice, however, the cross-sections are typically only piecewise constant, which necessitates the specification of interface conditions. Formulation of an appropriate set of these conditions (and corresponding boundary conditions) will be the subject of further investigation; possible approaches will be discussed in Sec. 5.4.3.

The fourth and third equation can be combined into:

$$\Sigma\psi^{(2)} = \nabla\cdot\left[\frac{3}{7\Sigma}\mathscr{S}\left(\nabla\otimes\psi^{(2)}\right) - \frac{6}{35\Sigma}\mathscr{S}\left(\mathbb{I}\otimes\nabla\cdot\psi^{(2)}\right)\right] - \frac{2}{5}\mathscr{S}\left(\nabla\otimes\psi^{(1)}\right) + \frac{2}{15}\mathbb{I}\nabla\cdot\psi^{(1)}.$$

Using identity (E.3), we obtain

$$\begin{aligned}\Sigma\psi^{(2)} = {} & \frac{3}{7\Sigma}\nabla\cdot\mathscr{S}\left(\nabla\otimes\psi^{(2)}\right) - \frac{6}{35\Sigma}\left[\frac{2}{3}\mathscr{S}\left(\nabla\otimes\nabla\cdot\psi^{(2)}\right) + \frac{1}{3}\mathscr{S}\left(\mathbb{I}\otimes\nabla\cdot\nabla\cdot\psi^{(2)}\right)\right] \\ & - \frac{2}{5}\mathscr{S}\left(\nabla\otimes\psi^{(1)}\right) + \frac{2}{15}\mathbb{I}\nabla\cdot\psi^{(1)}.\end{aligned}$$

Applying the identity (E.1) on the first term on right, we obtain the equation involving $\psi^{(2)}$ only through its divergence and laplacian:

$$\begin{aligned}\Sigma\psi^{(2)} = {} & \frac{2}{7\Sigma}\mathscr{S}\left(\nabla\otimes\nabla\cdot\psi^{(2)}\right) - \frac{6}{35\Sigma}\left[\frac{2}{3}\mathscr{S}\left(\nabla\otimes\nabla\cdot\psi^{(2)}\right) + \frac{1}{3}\mathscr{S}\left(\mathbb{I}\otimes\nabla\cdot\nabla\cdot\psi^{(2)}\right)\right] \\ & - \frac{2}{5}\mathscr{S}\left(\nabla\otimes\psi^{(1)}\right) + \frac{2}{15}\mathbb{I}\nabla\cdot\psi^{(1)} + \frac{1}{7\Sigma}\nabla^2\psi^{(2)}\end{aligned}$$

Using now eq. (5.30) for $\nabla\cdot\psi^{(2)}$ and a little bit of algebra, we obtain

$$\begin{aligned}\Sigma\psi^{(2)} = {} & -\frac{2}{35\Sigma}\mathscr{S}\left(\nabla\otimes\nabla\phi\right) + \frac{2}{105\Sigma}\mathbb{I}\nabla^2\phi \\ & - \frac{4}{7}\mathscr{S}\left(\nabla\otimes\psi^{(1)}\right) + \frac{4}{21}\mathbb{I}\nabla\cdot\psi^{(1)} \\ & + \frac{1}{7\Sigma}\nabla^2\psi^{(2)} \\ = {} & -\frac{2}{35\Sigma}\mathscr{D}\mathscr{S}\left(\nabla\otimes\phi\right) - \frac{4}{7}\mathscr{D}\mathscr{S}\left(\nabla\otimes\psi^{(1)}\right) + \frac{1}{7\Sigma}\nabla^2\psi^{(2)}.\end{aligned} \tag{5.33}$$

Equations (5.29), (5.30) and (5.33) form an equivalent MCP$_3$ system where the moment $\psi^{(3)}$ has been eliminated. Combining the first two equations yields

$$\nabla\cdot\left(-\frac{1}{3\Sigma}\nabla\phi - \frac{1}{\Sigma}\nabla\cdot\psi^{(2)}\right) + \Sigma_0\phi = q.$$

Since from eq. (5.33)

$$\begin{aligned}\Sigma\nabla\cdot\psi^{(2)} &= -\frac{4}{105\Sigma}\nabla^2\nabla\phi - \frac{2}{21}\nabla(\nabla\cdot\psi^{(1)}) - \frac{2}{7}\nabla^2\psi^{(1)} + \frac{1}{7\Sigma}\nabla^2\nabla\cdot\psi^{(2)} \\ &= -\frac{4}{105\Sigma}\nabla^2\nabla\phi - \frac{2}{21}\nabla(\nabla\cdot\psi^{(1)}) + \frac{1}{7\Sigma}\nabla^2\left(\nabla\cdot\psi^{(2)} - 2\Sigma\psi^{(1)}\right) \\ &= -\frac{4}{105\Sigma}\nabla^2\nabla\phi - \frac{2}{21}\nabla(\nabla\cdot\psi^{(1)}) - \frac{1}{7\Sigma}\nabla^2\left(\frac{1}{3}\nabla\phi + 3\Sigma\psi^{(1)}\right) \\ &= -\frac{3}{35\Sigma}\nabla^2\nabla\phi - \frac{2}{21}\nabla(\nabla\cdot\psi^{(1)}) - \frac{3}{7\Sigma}\nabla^2\psi^{(1)}\end{aligned}$$

(identity (E.2) has been used to obtain the first equality and eq. (5.30) to obtain the third) and hence

$$\Sigma\nabla\cdot\nabla\cdot\psi^{(2)} = -\frac{3}{35\Sigma}\nabla^4\phi - \frac{11}{21}\nabla^2(\nabla\cdot\psi^{(1)}) = -\frac{3}{35\Sigma}\nabla^4\phi - \frac{11}{21}\nabla^2(q - \Sigma_0\phi),$$

the MCP$_3$ equations imply that the scalar flux is governed by the following fourth-order PDE:

$$\frac{3}{35\Sigma^3}\nabla^4\phi - \left(\frac{11}{21}\frac{\Sigma_0}{\Sigma^2} + \frac{1}{3\Sigma}\right)\nabla^2\phi + \Sigma_0\phi = q - \frac{11}{21\Sigma^2}\nabla^2 q. \tag{5.34}$$

In the following subsection, we show that this equation may be recast into a system of diffusion-like equations equivalent with the SP$_3$ system (4.10).

5.4.2 Derivation of the SP$_3$-equivalent system

By multiplying with $\frac{21}{11}\Sigma$ and manipulating the terms, we put eq. (5.34) into the form

$$\frac{9}{55\Sigma^2}\nabla^4\phi + \frac{1}{\Sigma}\nabla^2(q - \Sigma_0\phi) - \frac{7}{11}\nabla^2\phi - \frac{21}{11}\Sigma(q - \Sigma_0\phi) = 0. \tag{5.35}$$

Let us define an auxiliary function θ such that

$$\frac{14}{11}\theta = -\frac{9}{55\Sigma^2}\nabla^2\phi - \frac{1}{\Sigma}(q - \Sigma_0\phi), \tag{5.36}$$

whence

$$\theta = -\frac{9}{70\Sigma^2}\nabla^2\phi - \frac{11}{14\Sigma}(q - \Sigma_0\phi), \tag{5.37}$$

and

$$-\frac{7}{11}\nabla^2\phi = \frac{490}{99}\Sigma^2\theta + \frac{35}{9}\Sigma(q - \Sigma_0\phi). \tag{5.38}$$

Comparing eq. (5.36) and first two terms in eq. (5.35), it follows that

$$-\frac{14}{11}\nabla^2\theta - \frac{7}{11}\nabla^2\phi - \frac{21}{11}\Sigma(q - \Sigma_0\phi) = 0.$$

After replacing the second term by eq. (5.38) and simple manipulations, this equation becomes a diffusion-like equation for θ:

$$-\frac{9}{14\Sigma}\nabla^2\theta + \frac{5}{2}\Sigma\theta - \Sigma_0\phi = -q. \tag{5.39}$$

The complementary diffusion-like equation for ϕ is obtained from (5.36):

$$-\frac{9}{55\Sigma}\nabla^2\phi + \Sigma_0\phi - \frac{14}{11}\Sigma\theta = q. \tag{5.40}$$

To facilitate the comparison between MCP$_3$ diffusion-like equations (5.40), (5.39) and SP$_3$ diffusion-like equations (4.10), we further multiply eq. (5.40) by 55/27 and eq. (5.39) by 14/27 to obtain

$$\begin{aligned} -\frac{1}{3\Sigma}\nabla^2\phi + \frac{55}{27}\Sigma_0\phi - \frac{70}{27}\Sigma\theta &= \frac{55}{27}q, \\ -\frac{1}{3\Sigma}\nabla^2\theta + \frac{35}{27}\Sigma\theta - \frac{14}{27}\Sigma_0\phi &= -\frac{14}{27}q. \end{aligned} \tag{5.41}$$

The SP$_3$ equations in the interior of a homogeneous region with isotropic scattering and sources read:

$$\begin{aligned} -\frac{1}{3\Sigma}\nabla^2\phi_0^s + \Sigma_0\phi_0^s - \frac{2}{3}\Sigma_0\phi_s^2 &= q, \\ -\frac{1}{7\Sigma}\nabla^2\phi_2^s + \left(\frac{4}{9}\Sigma_0 + \frac{5}{9}\Sigma\right)\phi_2^s - \frac{2}{3}\Sigma_0\phi_s^0 &= -\frac{2}{3}q. \end{aligned} \tag{5.42}$$

Recall that the scalar flux is obtained from the SP$_3$ solution by

$$\phi = \phi_0^s - \frac{2}{3}\phi_2^s \tag{5.43}$$

and formally corresponds to the zero-th moment of the 1D P$_3$ equations (4.6); from (4.2), we also have the formal correspondence between the second moments of the 1D P$_3$ and SP$_3$ equations:

$$\phi_2 = \frac{1}{3}\phi_2^s. \tag{5.44}$$

Substitution of (5.43) within (5.42) leads to

$$-\frac{1}{3\Sigma}\nabla^2\phi + \Sigma_0\phi = q, \tag{5.45}$$

$$-\frac{1}{7\Sigma}\nabla^2\phi_2^s + \frac{5}{9}\Sigma\phi_2^s - \frac{2}{3}\Sigma_0\phi = -\frac{2}{3}q. \tag{5.46}$$

Performing now the equivalent transformations

$$(5.45) \longrightarrow (5.45) - \frac{14}{9}\,(5.46),$$

$$(5.46) \longrightarrow \frac{7}{9}\,(5.46),$$

using (5.44) and putting $\theta = \phi_2$ transforms equations (5.45), (5.46) to the MCP$_3$ diffusion-like equations (5.41).

5.4.3 Direction for further research of interface conditions

In an infinite homogeneous medium, the two diffusion-like equations (5.45) and (5.46) with (5.37) may be solved for scalar flux ϕ equivalent to one that would obtained from the full set of 16 independent MCP$_3$ equations (5.29). Boundary and interface conditions need to be derived for solving problems in bounded and/or heterogeneous domains. Because of the formal equivalence with the SP$_3$ equations exhibited in the previous subsection, continuity of

$$\phi, \quad \theta, \quad \frac{1}{\Sigma}\frac{\partial\phi}{\partial\mathbf{n}}, \quad \frac{1}{\Sigma}\frac{\partial\theta}{\partial\mathbf{n}}$$

and boundary conditions analogous to those of the SP$_3$ model (Sec. 4.1.1) suggest themselves as a first approximation. Although these conditions let us find a unique solution of equations (5.45), (5.46), they are obviously not in general compatible with the MCP$_3$ model anymore. Generally, the conditions for ϕ and θ that ensure unique solvability of the coupled diffusion-like equations (5.45), (5.46) need to be obtained from appropriate conditions for the tensorial moments $\psi^{(n)}$ (which may be obtained as in the P$_N$ approximation using the formulation (5.28)).

A concrete set of interface conditions compatible with both the equations for angular flux moments and the diffusion-like equations of the form (5.45), (5.46)

has been presented in a rather puzzling paper [105] (see also commentary in [36, Sec. 5.2]). Starting from the even-parity NTE, the author "derives" equation (5.34) and equations (5.45), (5.46) as well as conditions for the even-parity tensorial moments and the unknowns of the diffusion-like equations – all on a half-page space. Even though the excessive brevity of the paper prevented further study and practical use of these equations and conditions, the equivalence of the equations to those systematically derived in previous subsections may motivate the search for the additional conditions in the same form as in [105].

6

Neutronics modules

As we have seen in Chap. 3, the various transport approximations lead to coupled systems of differential equations with finite element approximations formulated on product spaces. In order to implement these models into a practical solution method, it is thus necessary to appropriately deal with the coupling of variables.

In standard finite element assembling procedures, we would need to use the same underlying mesh $\mathcal{T}_h$ and approximation space $\mathcal{V}_{hp}$ for all components of the solution (which, given the large differences in behavior of the studied fields like group fluxes, directional fluxes, or angular flux moments, means a significant waste of computational resources) or use some form of data interpolation to evaluate integrals of type

$$\int_{\tau\in\mathcal{T}_h} \sigma_s^{gg'} u^g v^{g'} \,\mathrm{d}\mathbf{r} \tag{6.1}$$

(as appearing for instance in the multigroup diffusion approximation) when u^g and $v^{g'}$ do not "live" on the same mesh. To avoid this interpolation and errors associated with it, the C++ finite element library Hermes2D uses the original *multimesh assembling* approach [92].

To illustrate Hermes2D capabilities for efficiently solving coupled PDE systems, we will consider the general form of both the multigroup diffusion and SP_N problems as discussed in Sec. 4.2, i.e. the general form of Problem 4 (or, equivalently, 4′). As we are interested here in its finite element approximation, we start by writing the restriction of this general weak problem to the approximation subspace constructed in such a way that each component of the solution can be

approximated by different set of basis functions:

$$\mathbb{V}_{hp} \equiv \mathbb{V}_{hp}(\mathcal{D}) = \prod_{j=1}^{M} \mathcal{V}_j^{hp}(\mathcal{D}) \subset \mathbb{H}^1(\mathcal{D}),$$

where the component-specific spaces $\mathcal{V}_j^{hp}$ will be specified below.

Problem 5. *Given* $\mathrm{Q} = \mathrm{col}\,\{q_i\}_M \in [L^2(\mathcal{D})]^M$, *find* $\mathrm{U}_{hp} = \mathrm{col}\,\{u_j^{hp}\}_M \in \mathbb{V}_{hp}$ *such that*

$$\begin{aligned} a(\mathrm{U}_{hp}, \mathrm{V}) &= f(\mathrm{V}) \quad \forall \mathrm{V} \in \mathbb{V}_{hp}, \\ a(\mathrm{U}, \mathrm{V}) &:= \int_{\mathcal{D}} \big(\mathbf{D}\nabla \mathrm{U} : \nabla \mathrm{V} + \mathbf{C}\mathrm{U} \cdot \mathrm{V}\big)\, d\mathbf{r} + \int_{\partial\mathcal{D}} \gamma \mathbf{G}\mathrm{U} \cdot \mathrm{V}\, dS\,, \\ f(\mathrm{V}) &:= \int_{\mathcal{D}} \mathrm{Q} \cdot \mathrm{V}\, d\mathbf{r}\,, \quad \big(\mathbf{D}\nabla \mathrm{U}\big) : \nabla \mathrm{V} = \sum_{j=1}^{M} \sum_{\alpha=1}^{3} D_j \frac{\partial u_j}{\partial x_\alpha} \frac{\partial v_j}{\partial x_\alpha} \end{aligned} \tag{6.2}$$

where the diagonal matrix $\mathbf{D}$ *and matrices* $\mathbf{C}$ *and* $\gamma\mathbf{G}$ *are defined in* (4.9) *and Appendix C for the* $SP_{3,5,7}$ *models and in* (4.14) *for the multigroup diffusion* (SP_1) *model. As discussed in Sec. 4.2, eq.* (6.2) *can be put into the equivalent form*

$$\begin{array}{ccccccccccc} a_{11}(u_1^{hp}, v_1) & + & a_{12}(u_2^{hp}, v_1) & + & \cdots & + & a_{1M}(u_M^{hp}, v_1) & = & f_1(v_1), & \forall v_1 \in \mathcal{V}_1^{hp}, \\ a_{21}(u_1^{hp}, v_2) & + & a_{22}(u_2^{hp}, v_2) & + & \cdots & + & a_{2M}(u_M^{hp}, v_2) & = & f_2(v_2), & \forall v_2 \in \mathcal{V}_2^{hp}, \\ \vdots & & \vdots & & \ddots & & \vdots & = & \vdots & \\ a_{M1}(u_1^{hp}, v_M) & + & a_{M2}(u_2^{hp}, v_M) & + & \cdots & + & a_{MM}(u_M^{hp}, v_M) & = & f_M(v_M), & \forall v_M \in \mathcal{V}_M^{hp} \end{array}$$

where for $i, j = 1, 2, \ldots, M$,

$$a_{ij}(u, v) := \int_{\mathcal{D}} \big(D_{ij}\nabla u \nabla v + C_{ij} uv\big)\, d\mathbf{r} + \int_{\partial\mathcal{D}} \gamma G_{ij} uv\, dS\,, \quad f_i(v) := \int_{\mathcal{D}} q_i v\, d\mathbf{r}\,.$$

6.1 Multimesh hp-FEM

Let us define a set of meshes

$$\mathcal{T}_j = \{\tau_j\}, \quad \overline{\mathcal{D}} = \bigcup_{\tau_j \in \mathcal{T}_j} \overline{\tau}_j.$$

We assume that the meshes originated from a common *master mesh* $\mathcal{T}^m$ by successive refinements (independent for each mesh), but are otherwise arbitrary. The corresponding approximation spaces are given by

$$\mathcal{V}_j^{hp} = \left\{ v_j^{hp} \in C^0(\mathcal{D}) : v_j^{hp}\Big|_{\tau_j} \circ \mathfrak{r}_j \in \mathcal{P}_p(\hat{\tau}),\ \tau_j \in \mathcal{T}_j \right\} \tag{6.3}$$

where $\hat{\tau}$ is the reference element (unit square or right triangle in Hermes2D), $\mathfrak{r}_j : \tau_j \to \hat{\tau}$ maps the physical element in j-th mesh to the reference element and $\mathcal{P}_p(\hat{\tau})$ is the space of polynomials of degree up to p (tensor product polynomials in case of quadrilateral τ). We will refer to the highest degree p of the polynomial contained in the approximation space as to the approximation order of that space. In Hermes2D, the meshes can contain both triangular and quadrilateral elements. In order to prevent condition numbers of the standard finite element stiffness matrices from blowing up with increasing polynomial degree, hierarchical shape functions are used to construct the approximation spaces (6.3) in preference to the usual Lagrangian nodal basis ([107, Sec. 2.5.3]). In the case of a quadrilateral mesh, for instance, $\mathcal{V}_j^{hp}$ is composed of functions that are on the unit square $\hat{\tau}$

(a) $p \geq 1$: bilinear functions with value 1 in exactly one vertex of $\hat{\tau}$ and 0 in the remaining vertices,

(b) $p \geq 2$: 2D polynomials with nonzero values on exactly one edge of $\hat{\tau}$ and vanishing on the remaining edges,

(c) $p \geq 4$: 2D polynomials with nonzero values in the interior of $\hat{\tau}$ and vanishing on $\partial\hat{\tau}$.

(see Fig. 6.1).

Note that the local approximation subspaces associated with each quadrilateral element might possess different approximation orders in the lateral and

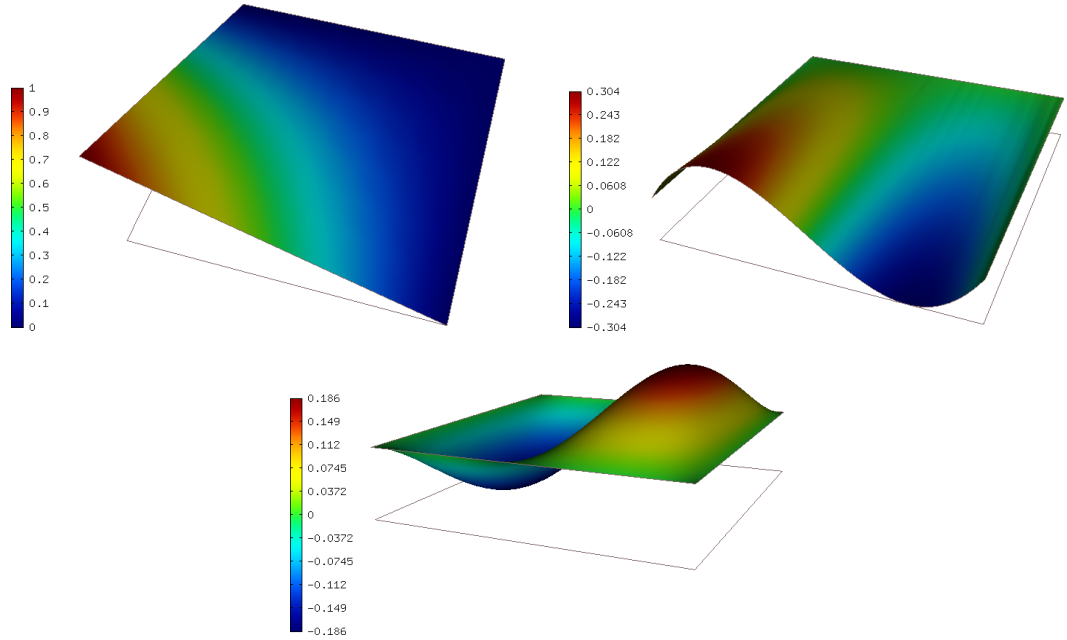


Figure 6.1: Shape functions of type (a), (b), (c).

longitudinal direction (as they are defined by a tensor product of 1D polynomials in these two directions). Moreover, approximation orders may also vary from element to element, i.e. for each j, $\mathcal{V}_j^{hp}$ is decomposed into

$$\mathcal{V}_{j,\tau}^{hp} = \{v_j^{hp} \circ \mathfrak{r}_j \in \mathcal{P}_{p_j}(\hat{\tau})\}, \quad \tau_j \in \mathcal{T}_j, \quad p_j = p(\tau_j)$$

the local approximation subspaces constructed so that the minimum rule for H^1-conforming approximations ([90, §3.5.5]) – namely that the approximation order of these subspaces coincides with the approximation order in element interiors (thus constraining the polynomial degree of the edge shape functions during assembling) – is satisfied. We will henceforth identify an element with the local approximation subspace constructed on top of it, so that we can speak of element orders, etc.

Lastly, hanging nodes (mesh vertices in edge interiors) are also allowed in $\mathcal{T}_j$ for greater flexibility of mesh adaptivity (and H^1 conformity recovered by additional constraints on the edge shape functions as described in [90, §3.6] and [89]).

6.1.1 Multimesh assembling

We are now ready to describe the basic principle of the multimesh assembling algorithm. This has been nicely done in the original paper [92] but we include the brief description here as well (closely following [92]) as it is directly related to one of author's own contributions to Hermes2D described in Sec. 6.1.2. For illustration, we consider $M = 3$, i.e. three different meshes.

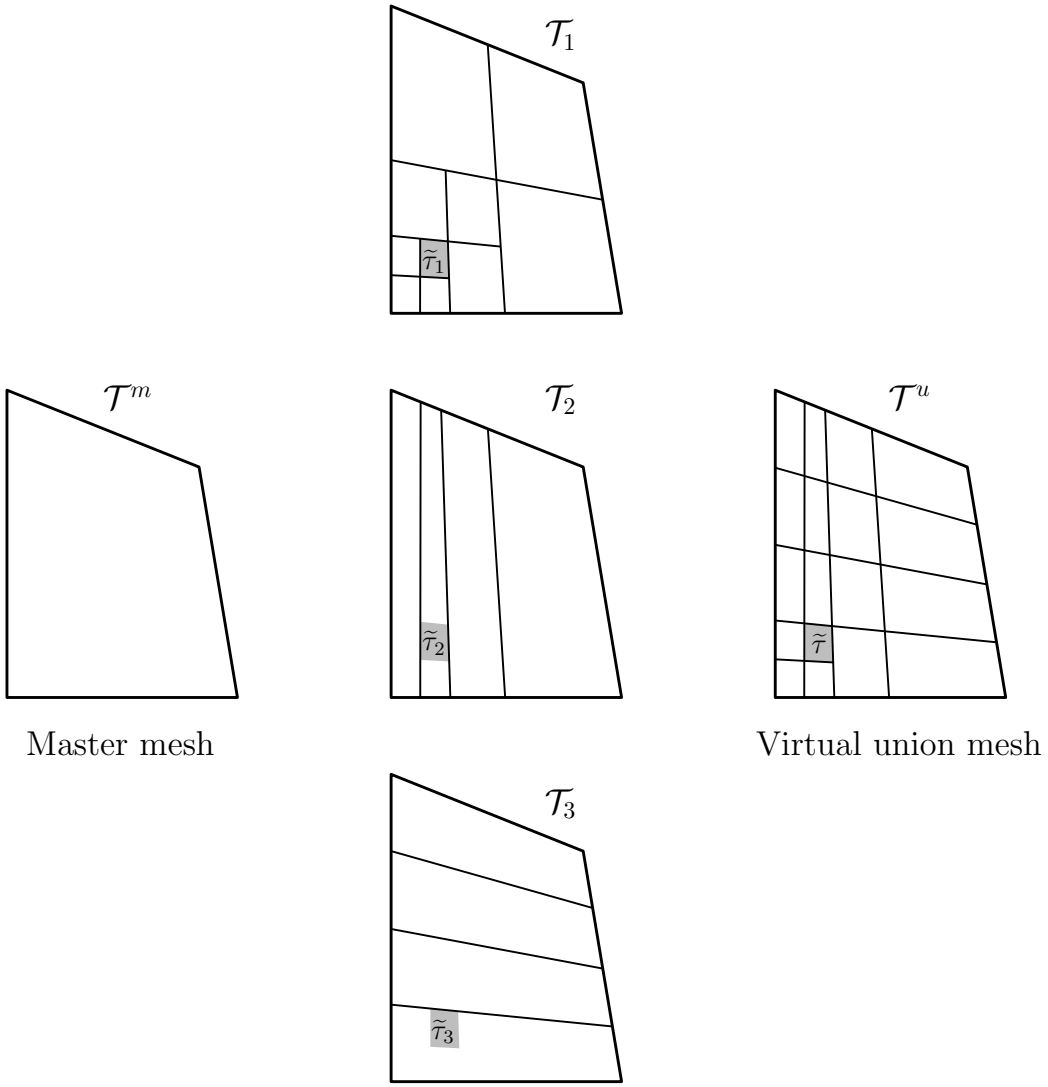


Figure 6.2: One state of the multimesh assembling algorithm. Note that the sub-element mapping $\mathfrak{s}_1$ is identity.

In order to assemble mixed integrals of type (6.1), the assembling procedure works with a geometrical union $\mathcal{T}^u$ of all meshes $\mathcal{T}_1$, $\mathcal{T}_2$, $\mathcal{T}_3$. This union is never explicitly created in memory. Rather, its virtual elements are traversed by the usual element-wise assembling loop. Let $\widetilde{\tau} \in \mathcal{T}^u$ be the currently visited virtual element. As all meshes $\mathcal{T}_j$ originate from a common master mesh, there is for each j exactly one element $\tau_j \in \mathcal{T}_j$ such that $\widetilde{\tau}_j \subset \tau_j$ is the *sub-element* of τ_j (which is called in this context the *active element* on $\mathcal{T}_j$) corresponding to $\widetilde{\tau}$.

To evaluate the integrals comprising the weak formulation, each sub-element needs to be transformed to the reference element where the appropriate quadrature points are defined. While for the active element on $\mathcal{T}_j$, the referece mapping $\mathfrak{r}_j(\tau_j) = \hat{\tau}$ as required, it transforms the sub-element to a subset $\mathfrak{r}_j(\widetilde{\tau}_j) \subset \hat{\tau}$. Thus another mapping $\mathfrak{s}_j : \hat{\tau} \to \hat{\tau}$ is introduced (the *sub-element mapping*) such that $\mathfrak{s}_j(\mathfrak{r}_j(\widetilde{\tau}_j)) = \hat{\tau}$. These two mappings allow Hermes2D to evaluate all integrals in the weak forms by using elements only on the mesh on which the integrands live, thus incurring no further error beyond the inevitable one of numerical integration.

6.1.2 Discontinuous Galerkin assembling

As a joint effort of the author of this thesis and Lukáš Korous, at this moment (November 2014) the main developer of the Hermes2D project, Hermes2D has been enabled for discontinuous Galerkin approximation of variational problems. This involved the extension of the assembly procedure to perform surface integration over all edges of all elements for user-defined discontinuous Galerkin bilinear and linear forms and also exposing for these forms the access to shape functions and geometrical information of both elements sharing a common interface (thus allowing the user to define interface operators like $[\![\cdot]\!]$ and $\langle\cdot\rangle$ introduced in Sec. 3.5.2.1). Because of the possibility of hanging nodes in the mesh (essentially needed for efficient mesh refinement), the actual integration of these forms along element edges is non-trivial as matching points from both sides of the edge need to be correctly determined. This functionality has been implemented in the class NeighborSearch based on the same fundamental idea underlying the multimesh assembling (namely that of sub-element mappings).

The `NeighborSearch` class characterizes a neighborhood of a given edge in terms of adjacent elements and provides methods for getting limit values of discontinuous functions from both sides of the edge. Each instance of the class is connected to a mesh and its active element. The current active element becomes the *central element* of the neighborhood and all adjacent elements the *neighbors*. In order to search for the neighboring elements, one selects a particular edge of the central element and calls a function that enumerates the neighbors and fills in the array of sub-element mappings (transformations) necessary for getting function values at matching quadrature points from both sides of the selected (active) edge. The actual procedure depends on the relative size of the central element with respect to the neighbor element(s) across the active edge:

- If active edge is shared by two elements of same size, then the neighboring element is identified directly and no additional transformations are needed to obtain values of any given function from either side of active edge.

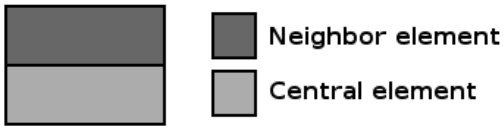


Figure 6.3: Neighbor search – "no transformation" case.

- If the neighbor element is bigger than the central element, then we "go up" on the central element, until we find its parent that has the same size as the neighbor. We keep track of the visited intermediate parents and after the final one has been found (in the ultimate case an element of the master mesh), we use them in reverse order to fill in the sub-element mapping array. These transformations will be applied to integration points used when integrating a function on the neighboring (bigger) element in order to obtain values at points matching those from the central element's side.

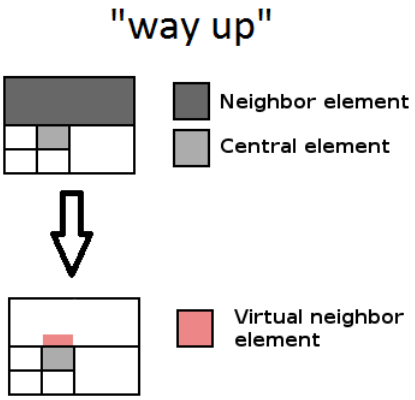


Figure 6.4: Neighbor search – "way up" case.

- If the neighbor element is smaller than the central element, it means that it is one of several neighbors across the active edge. Hence, we "go down" in the central element in order to find a (virtual) sub-element matching the currently processed neighbor and store the corresponding transformations in the neighbor's row of the (two-dimensional) array of sub-element mappings. This way, we obtain for each neighbor a set of transformations which will be applied on the central element to transform integration points to the correct sub-element matching the neighbor.

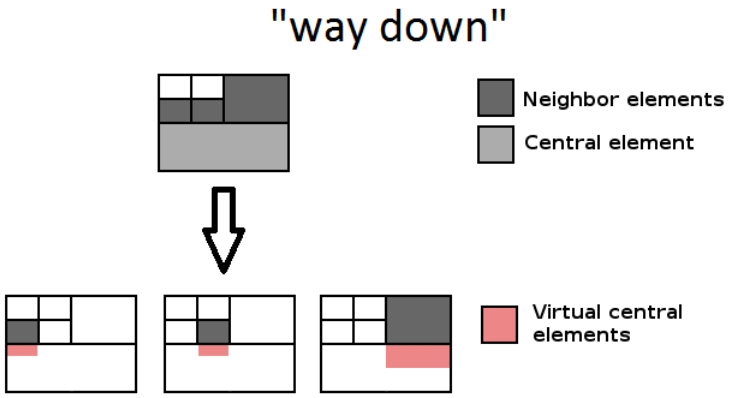


Figure 6.5: Neighbor search – "way down" case.

If only an external function is supposed to be discontinuous across the active edge (e.g. in the case of assembling linear forms), an appropriate `DiscontinuousFunc` object is created for such a function and is exposed to the

user who can use it to retrieve actual values at the matching integration points from both sides of the active segment.

If also the trial and test functions need to be considered discontinuous (e.g when assembling DG bilinear forms), the local approximation bases (called *shapesets* in Hermes2D) on the central element and on the current neighbor element are extended by zero to the whole neighborhood of the two elements. The so-called *extended shapeset* thus created is queried during assembling for the count of all contained extended shape functions and their global DOF (degree of freedom) numbers. Assembling is done over all these extended shape functions – the currently processed trial and test functions is exposed to the user again as `DiscontinuousFunc` objects with the shape functions' values (and derivatives) at integration points at both sides of active edge.

This procedure is limited by the requirement that the number of integration points at both sides of active edge is the same. This is enforced artificially during assembling by performing the integrations of DG interface forms using a quadrature of the maximal supported order. Note that this does not restricts the approximation spaces in any way – it only pertains to the numerical integration.

6.2 *hp*-adaptivity

The goal of the *hp*-adaptivity process is to combine spatial subdivision of selected elements (h-refinement) with local increase of their approximation order (p-refinement) so that the total available number of degrees of freedom at given stage of the process is utilized most efficiently (i.e., relatively big number of low-order elements is used in regions with highly oscillating solution while smaller number of high-order elements in regions with smoother solution behavior). As there are many possible combinations of h and p refinements of given element, one number per element provided by traditional a-posteriori error estimates for FEM is insufficient to guide the *hp* adaptivity. To circumvent this issue, local error estimation in Hermes2D is based on comparing two solutions of different approximation orders – a robust technique borrowed from the field of numerical solution of ordinary differential equations. In particular, it allows to determine the whole shape of the approximation error $e = u - u_{hp}$ over each element and use it to

determine the best refinement candidate that decreases the total approximation error for the lowest number of added DOF.

For illustration, let us consider the general approximate problem 5, coming from the corresponding exact variational problem with solution $\mathrm{U} \in \mathbb{H}^1(\mathcal{D})$. Note that the approximation error is a vector-valued function

$$\mathrm{E}_{hp} = \mathrm{U}_{hp} - \mathrm{U};$$

the total error can be measured by

$$\|\mathrm{E}_{hp}\|^2 = a(\mathrm{E}_{hp}, \mathrm{E}_{hp}) = \sum_{i=1}^{M}\sum_{j=1}^{M} a_{ij}(\mathrm{e}_j^{hp}, \mathrm{e}_i^{hp}) = \sum_{i=1}^{M} \|\mathrm{e}_i^{hp}\|^2, \tag{6.4}$$

which in the case of symmetric coercive bilinear form a (single-group diffusion/SP_N models) is the true energy norm associated with the problem. In the non-symmetric case, (6.4) can still be used to measure the approximation error and guide the adaptivity. This approach takes into account the coupling of the fields (group-to-group coupling induced by scattering and/or the coupling between the SP_N moments). Alternatively, one may use the simpler $\mathbb{H}^1(\mathcal{D})$ norm – then

$$a(\mathrm{U}, \mathrm{V}) = (\mathrm{U}, \mathrm{V})_{\mathbb{H}^1(\mathcal{D})}, \quad a_{ij}(u, v) = (u, v)_{H^1(\mathcal{D})}\delta_{ij}$$

and the $\mathbb{H}^1(\mathcal{D})$ norm of the total error is given by

$$\|\mathrm{E}_{hp}\|_1^2 = \sum_{i=1}^{M} \|\mathrm{e}_i^{hp}\|_1^2 = \sum_{i=1}^{M} (\mathrm{e}_i^{hp}, \mathrm{e}_i^{hp})_{H^1(\mathcal{D})}.$$

To guide the adaptivity process, the error is estimated by replacing the unknown exact solution U by a *reference solution* $\mathrm{U}_{\mathrm{ref}}$ approximating U using a greater number of degrees of freedom than U_{hp}. The adaptation algorithm consists of the following steps:

1. Given a set of spaces $\{\mathcal{V}_j^{hp}\}_M$, a uniformly refined set $\{\mathcal{V}_j^{h/2,p+1}\}_M$ is first created and Problem 5 solved on these spaces, yielding $\mathrm{U}_{h/2,p+1} =: \mathrm{U}_{\mathrm{ref}}$.

2. The coarse component of approximation pair $(\mathrm{U}_{hp}, \mathrm{U}_{\mathrm{ref}})$ is then obtained as

$$\mathrm{U}_{hp} = \Pi_{hp}\mathrm{U}_{h/2,p+1}.$$

   We still denote the error function defined via this approximation pair as E_{hp}:

$$\mathrm{E}_{hp} = \mathrm{U}_{hp} - \mathrm{U}_{\mathrm{ref}} = \Pi_{hp}\mathrm{U}_{h/2,p+1} - \mathrm{U}_{h/2,p+1}.$$

3. The local contribution of each element $\tau_i \in \mathcal{T}_i$ to the i-th component of the total error

$$\|\mathrm{e}_i^{hp}\|^2 = \sum_{j=1}^{M} a_{ij}(\mathrm{e}_j^{hp}, \mathrm{e}_i^{hp}), \quad i = 1, 2, \ldots, M$$

   can now be computed during the multimesh assembly of the involved bilinear forms. Let us denote this contribution by $\eta_{\tau,i}^{hp}$.

4. Elements contributing the most to the global error are marked for refinement (using a user-specified threshold). Note that elements of *all* meshes are considered, so that every element of every mesh in $\{\mathcal{T}_i\}_M$ is compared with all others.

5. For each element marked for refinement, a multitude of refinements is tested, so that $\tau_i \in \mathcal{T}_i$ in space $\mathcal{V}_i^{hp}$ gives rise to several candidate spaces $\mathcal{V}_i^{h'p'}$ (83 possibilities are tested for triangular elements as illustrated below and even more possibilities arise in the case of quadrilateral elements by considering anisotropic refinements in longitudinal and lateral directions – for more details, we refer to [89, 91]). Error contribution $\eta_{\tau,i}^{h'p'}$ of each candidate is computed from the difference

$$\Pi_{h'p'}\mathrm{U}_{h/2,p+1} - \mathrm{U}_{h/2,p+1}.$$

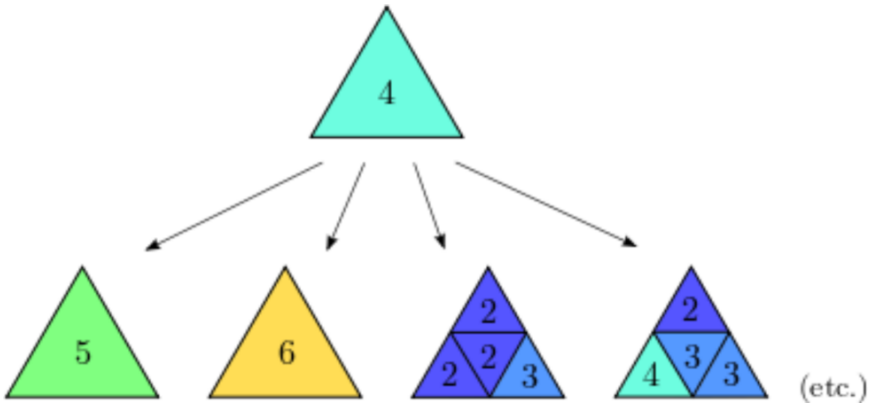


Figure 6.6: Refinement candidates of a triangular element (number signifies the approximation order). From `http://hpfem.org/wp-content/uploads/doc-web/doc-tutorial/src/hermes2d/D-adaptivity/intro-1.html`

6. The refinement candidate that leads to the largest decrease of error for the associated increase of the number of DOF

$$\frac{\log\left(\eta_{\tau,i}^{hp}\right) - \log\left(\eta_{\tau,i}^{h'p'}\right)}{\left(N_{\mathrm{DOF}}^{h'p'} - N_{\mathrm{DOF}}^{hp}\right)^{\alpha}}.$$

where α is a customizable convergence exponent (set to 1 by default), is finally selected and τ_i is refined accordingly, ultimately leading to a set of adapted spaces for another iteration. The whole process is terminated when the total error norm estimate $\|\mathrm{E}_{hp}\|$ decreases below specified threshold.

These steps represent a straightforward generalization of the so called projection-based *hp*-adaptivity [35, 107] to multimesh setting. This approach is sufficiently general so that it can be employed in solution of any type of PDE or their system. Also, pure *h*- or *p*-adaptivity can be realized by the same algorithm. For an *h*-adaptive solution of simple elliptic equations, however, there exists a big number of simpler and potentially more efficient approaches based on standard a-posteriori error estimates [54]. One based on the estimate of the energy norm of the error by the sum of the element residual and the jump of solution at element boundaries (computed using the actual approximation U_{hp}) [66] has been also implemented into Hermes2D using the newly added discontinuous Galerkin capability (needed for evaluating the jump terms) and will be used in example Sec. 6.3.1.2.

6.2.1 Error estimator for SP_N based on the scalar flux

As we are typically most interested in the scalar flux (rather than the SP_N auxiliary fluxes), we may use (4.5) to estimate the error in scalar flux approximation for the SP_N model (where we recall that N is taken as odd integer):

$$\begin{aligned}\|\mathrm{E}_{hp}^{\phi}\|_1^2 &= \left\| \sum_{n=0}^{(N-1)/2} F_n\big(\phi_{2n}^s - \phi_{2n}^{s,\mathrm{ref}}\big) \right\|_1^2 \leq \left(\sum_{n=0}^{(N-1)/2} \left\| F_n\big(\phi_{2n}^s - \phi_{2n}^{s,\mathrm{ref}}\big) \right\|_1 \right)^2 \\ &= \left(\sum_{n=0}^{(N-1)/2} \|F_n e_{2n}^s\|_1 \right)^2 = \sum_{m=0}^{(N-1)/2} \sum_{n=0}^{(N-1)/2} F_n F_m \|e_{2n}^s\|_1 \|e_{2m}^s\|_1 \\ &= \sum_{m=0}^{(N-1)/2} \sum_{n=0}^{(N-1)/2} a_{mn}^{\phi}(e_{2n}^s, e_{2m}^s)\end{aligned} \tag{6.5}$$

where $e_{2n}^s = \big(\phi_{2n}^s - \phi_{2n}^{s,\mathrm{ref}}\big)$ and $a_{mn}^{\phi}(u,v) = F_n F_m \|u\|_1 \|v\|_1$. This estimate will be used in the SP_N examples in Sec. 6.3.1 to evaluate local element error contributions as described in step 3 of the algorithm above.

6.3 Neutronics modules and examples

A multigroup SP_N/diffusion framework has been developed on top of Hermes2D during the work on the thesis. It implements all the basic building blocks of the weak formulation of Problem 5 for both models It also allows relatively simple specification of typical transport benchmarks, with automatic conversion between various possible sets of specified cross-sections (basically implementing relations like (2.14) in the multigroup approximation) and calculating and visualizing basic quantities of interest like ϕ, J and reaction rates. The SP_N equations have been implemented and tested up to $N = 9$, using the Mathematica script referred to in App. C, which allows simple extension to higher orders.

For solving the criticality eigenvalue problems of type (3.5), the (inexact) inverse iteration has been implemented according to paper [50]. It supports both fixed and Rayleigh quotient shifting and exact and inexact solves (with fixed or

decreasing tolerance during the course of the eigenvalue iteration) of the associated linear system in the i-th iteration:

$$(\mathbf{A} - s^{(i)}\mathbf{B})\mathbf{x}^{(i+1)} = \mathbf{B}\mathbf{x}^{(i)},$$

where $s^{(i)}$ is the actual shift value. Combination of both shift types is possible as it may be generally advantageous to perform several iterations with fixed shift (possibly zero, when the iteration reduces to the simple power iteration) and then to switch to variable shifting to speed-up the convergence rate.

In addition, the S_N method has been also implemented to test the accuracy of the more approximate SP_N and diffusion models. This module uses and extends the neutronics classes developed for the SP_N module (material data specification, visualization, the class that encapsulates fixed-point iteration for both the eigenvalue calculation and simple source iteration (3.39), etc.) and implements the discontinuous Galerkin spatial discretization as described in Sec. 3.5.1, using the multimesh DG assembler outlined in Sec. 6.1.2. The ordinates set described in Sec. 3.4.4 has been implemented in the module (note that for solving 2D problems to which Hermes2D is restricted, only $M/2 = N(N+2)/2$ ordinates in the upper hemisphere $\Omega_z > 0$ need to be considered in constructing the S_N system). As such, this ordinates set allows to accurately represent reflective conditions at boundaries aligned with axes of the Cartesian coordinate system.

For solving the multigroup diffusion problems, this framework is available in the master branch of current Hermes2D release. The SP_N version with various examples presented below is available from a local repository of the author and is based on an older version of Hermes:
`https://github.com/mhanus/hermes/tree/neutronics-master`.
The development version mainly for testing the S_N module (including additional S_N examples) is available from
`https://github.com/mhanus/hermes/tree/SN-adaptive`.

6.3.1 SP$_N$ and diffusion examples

The following examples show the behavior of the SP$_N$ approximation and some features of Hermes2D and its new neutronics modules. The use of adaptivity kept the sizes of these problems reasonably small so that a sparse direct solver UMFPACK [32] could be used for their efficient solution. The calculations were performed on the Toshiba Portege M750 laptop with Intel Core 2 Duo P8400 2.26 GHz processor (3 MB L2 cache) and 4 GB DDR2 SDRAM. The operating system was Ubuntu 10.10 64bit.

6.3.1.1 Benchmark IAEA EIR-1

This classical neutronics benchmark of the International Atomic Energy Agency models an arrangement of four homogeneous absorbing and scattering regions with dimensions 30×25 cm placed at the center of a rectangular pool of total width 96 cm and height 86 cm with vacuum outer boundary conditions. Isotropic external neutron source is uniformly distributed in two of the regions as seen from Fig. 6.7 and Tab. 6.1 listing the material properties of the regions.

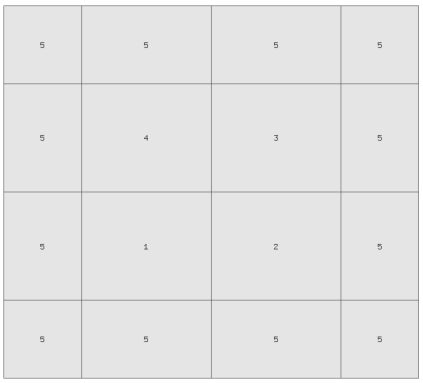


Figure 6.7: Initial mesh for the IAEA EIR-1 benchmark.

Figures 6.8 show the shape of the SP$_5$ solution. It consists of scalar flux and even-order moments ϕ_2^s, ϕ_4^s (SP$_N$ fluxes) as well as odd-order vector moments (SP$_N$ currents) obtained from the even order moments by eq. (4.4) (with derivative replaced by gradient). Notice the decreasing magnitude of the higher-order moments – the shape of the scalar flux is largely determined by the zero-th

Region #	σ_t [cm^{-1}]	σ_s [cm^{-1}]	q [cm^{-3}·s]
1	0.60	0.53	1.0
2	0.48	0.20	0.0
3	0.70	0.66	1.0
4	0.65	0.50	0.0
5	0.90	0.89	0.0

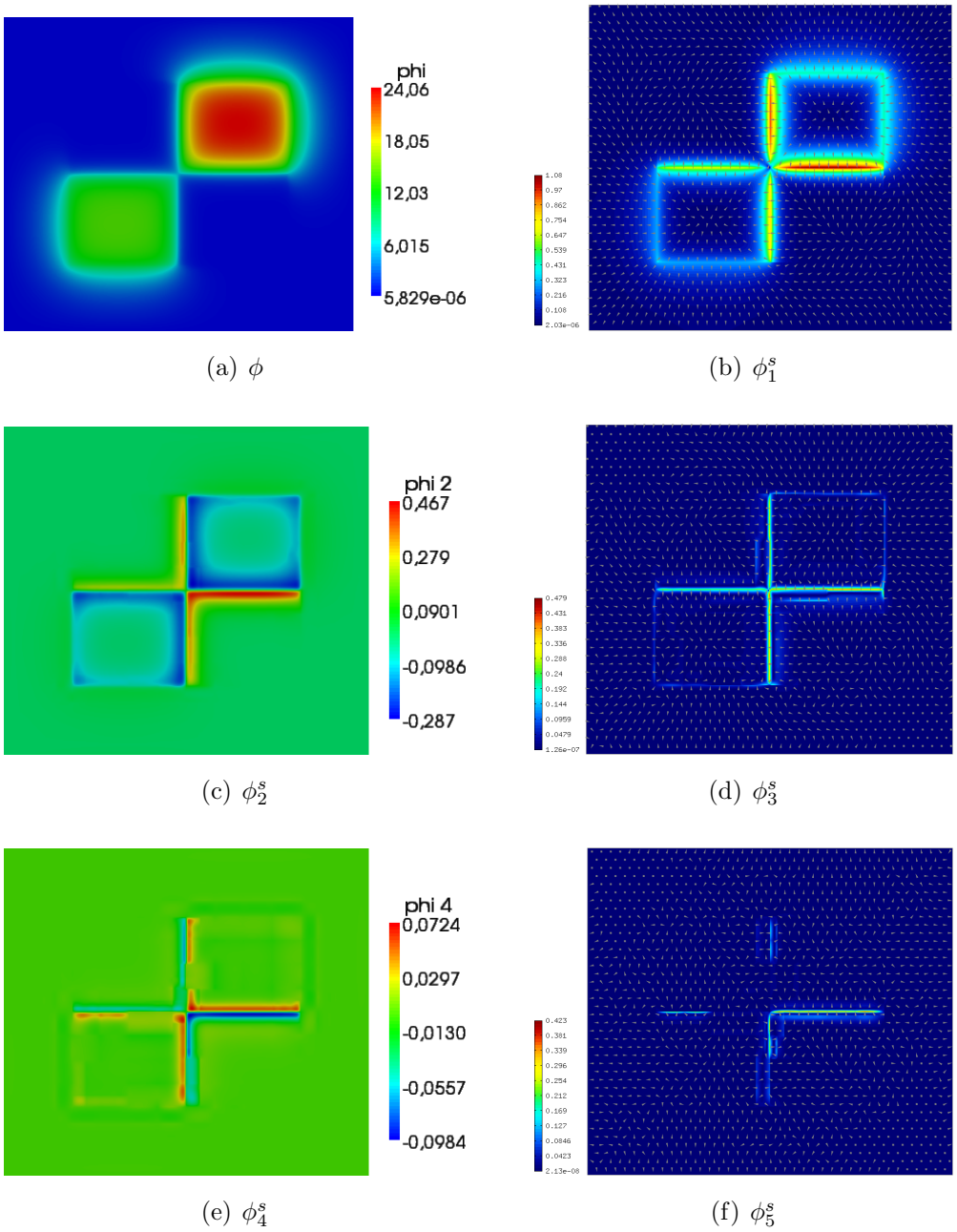
Table 6.1: Material properties for the IAEA EIR-1 benchmark.

SP_5 moment ϕ_0^s, while the higher-order moments provide transport corrections mostly visible at region interfaces. Also recall from (4.5) that the contributions of higher-order moments to the scalar flux are further weighted by factors < 1. Nevertheless, even these small corrections can have important impact on the overall results, as shown in Tab. 6.2, where average scalar fluxes in regions 1 to 5 are compared with the S_8 results reported by [25]. Note that relatively large homogeneous regions in this example (typical for whole-core reactor calculations) make it fit well into the asymptotic region of validity of the SP_N approximation.

Region #	SP_1	SP_3	SP_5	SP_7
1	0.81	0.12	0.06	0.05
2	5.23	0.80	0.31	0.26
3	0.90	0.14	0.07	0.07
4	3.86	0.61	0.39	0.37
5	0.84	0.06	0.06	0.08
t_{CPU} [s]	4.1	16.7	42.5	66.7

Table 6.2: IAEA EIR-1 benchmark: rel. errors [%] of average scalar flux in regions $i = 1, \ldots, 5$ w.r.t. S_8.

The solution was spatially converged using the *hp*-adaptivity of Hermes2D with convergence criterion $\|\mathrm{E}_{hp}^{\phi}\|_1 < 0.5\%$ (chosen as a reasonable value from the engineering point of view). Figure Fig. 6.11 shows the convergence curves for the SP_N fluxes (labeled "pseudo-fluxes" in the figure), scalar fluxes and the

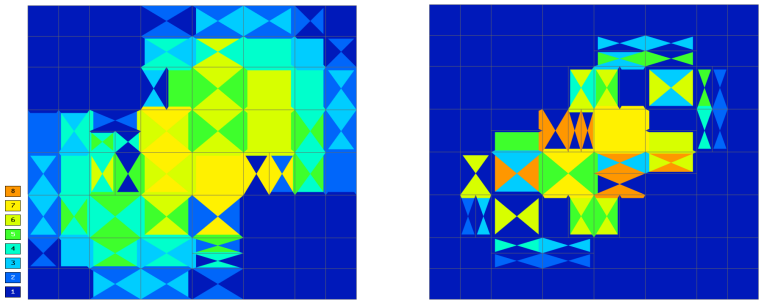


(a) ϕ (b) ϕ_1^s

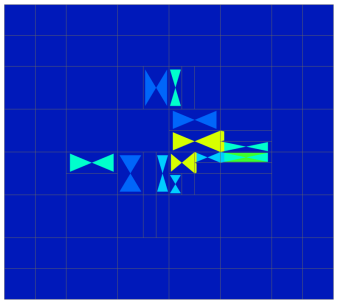
(c) ϕ_2^s (d) ϕ_3^s

(e) ϕ_4^s (f) ϕ_5^s

Figure 6.8: Solution of the IAEA EIR-1 benchmark (SP_5 approximation). Even order moments visualized using Paraview, odd order moments using Hermes2D internal functions.



(a) ϕ (b) ϕ_2^s



(c) ϕ_4^s

Figure 6.9: Approximation spaces for the IAEA EIR-1 benchmark (SP_5).

region-average scalar fluxes. Note how the scalar flux error estimate stays closely below the combined estimate of the SP_N fluxes error, verifying (6.5).

Figure 6.9 shows the distribution of final mesh sizes and local approximation orders and emphasize the utility of multimesh approximation and anisotropic adaptivity. As a legend for these and later approximation order figures, consider the element in Fig. 6.10. The colors encode that a tensor product of 1D polyno-

Figure 6.10: Legend for the approximation order figures.

mials of degree up to 3 and 8 in the horizontal and vertical direction, respectively,

is used to define the local approximation space for this element. Edge shape functions are constrained by edge shape functions of the adjacent elements, so as to satisfy the H^1 conformity constraints ([90, §3.5.5]).

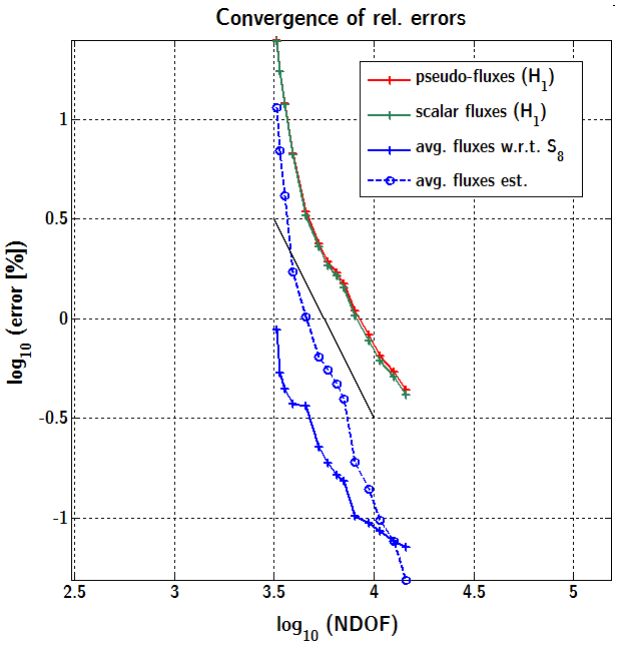


Figure 6.11: Adaptivity convergence curves for the IAEA EIR-1 benchmark (black line corresponds to fourth-order convergence) – SP_5.

6.3.1.2 The 7×7 PWR assembly example (Stankovski benchmark)

This example from [80, Sec. 6.4.2] represents a one-group calculation of neutron flux distribution in a small-scale model of a typical pressurized water reactor fuel assembly. The isotropic fixed source of $1.0\,\mathrm{cm}^{-3}$·sec is specified in all regions marked as `cell#` in Fig. 6.12 so as to represent neutrons scattered into the first group during their slowing-down in the moderator regions. Each moderator region has $\sigma_t = 1.250\,\mathrm{cm}^{-1}$, $\sigma_s = 1.242\,\mathrm{cm}^{-1}$, while each fuel region marked as `pin#` in Fig. 6.12 with the exception of `pin0` has $\sigma_t = 0.625\,\mathrm{cm}^{-1}$, $\sigma_s = 0.355\,\mathrm{cm}^{-1}$. Purely absorbing region `pin0` with $\sigma_t = 14\,\mathrm{cm}^{-1}$ corresponds to a control rod. The width of each square cell is 0.45 cm and fuel pin radius is 0.45 cm and vacuum boundary is assumed at the right side while reflective boundary at the remaining

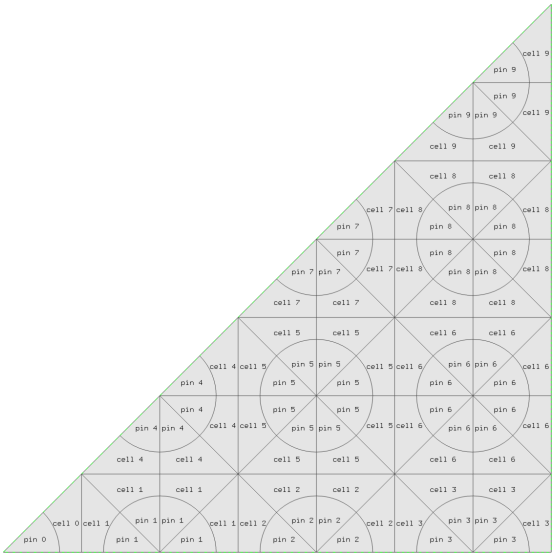


Figure 6.12: Geometry and initial mesh of the Stankovski benchmark.

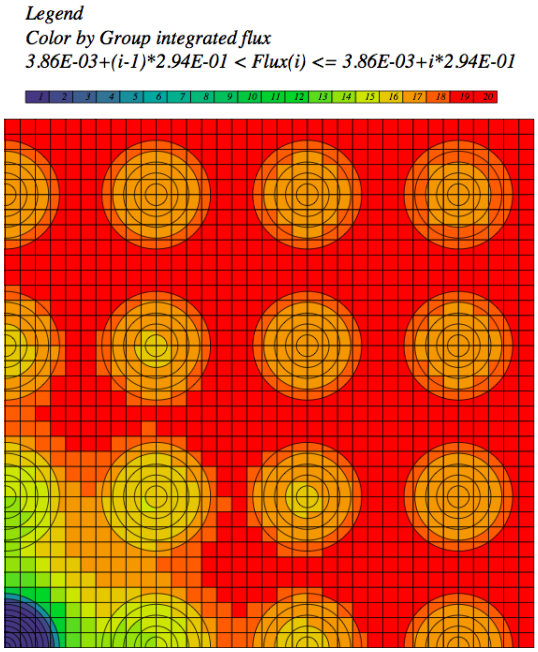
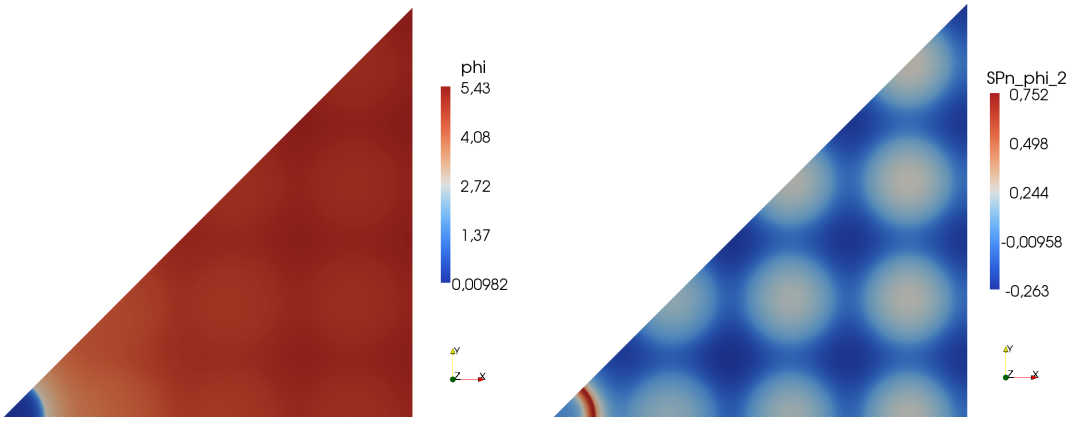


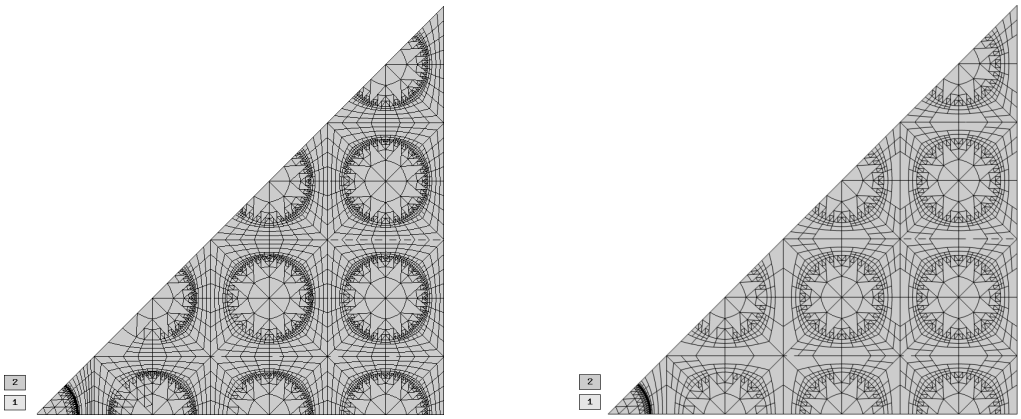
Figure 6.13: Stankovski benchmark – reference solution by DRAGON (method of characteristics, 2048 regions, quadrature order 20, integration lines spacing $100\,\mathrm{cm}^{-1}$); only a first quadrant is shown (extends by reflectional symmetry). Values span the range $[3.86 \times 10^{-3}, 5.88386]$.

sides. The ability of Hermes2D to use finite elements with curved boundaries with known non-uniform rational B-spline (NURBS) parametrization [90] has been utilized to treat the circular fuel pin boundary exactly.



(a) ϕ (b) ϕ_2^s

Figure 6.14: Solution of the Stankovski benchmark (SP_3).



(a) ϕ_0^s (b) ϕ_2^s

Figure 6.15: Approximation spaces for the Stankovski benchmark (SP_3).

Figures 6.14 and 6.15 show the SP_3 solution with the underlying meshes obtained from the simple h-adaptivity based on the error estimate of Kelly et al. as outlined in the last paragraph of Sec. 6.2. For comparison, the SP_1 (diffusion), SP_5 and S_8 solutions are also shown in figures 6.17, 6.18 and 6.19, respectively

(3× globally refined initial mesh from Fig. 6.12 and piecewise linear DG(1) elements have been used for the S_8 calculation). Note that scalar flux is displayed instead of ϕ_0^s in the SP_N figures to facilitate the comparison.

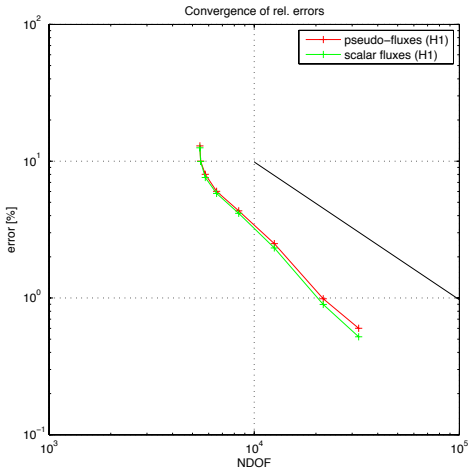


Figure 6.16: Adaptivity convergence curves for the Stankovski benchmark (black line corresponds to second-order convergence) – SP_5 approximation, bi-quadratic finite elements.

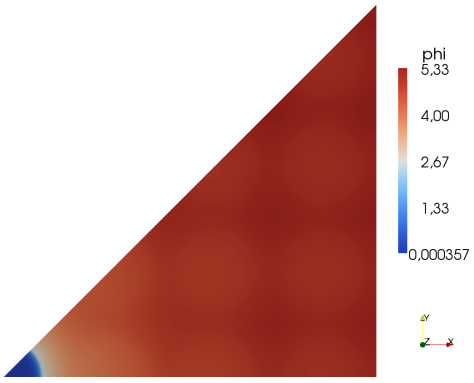
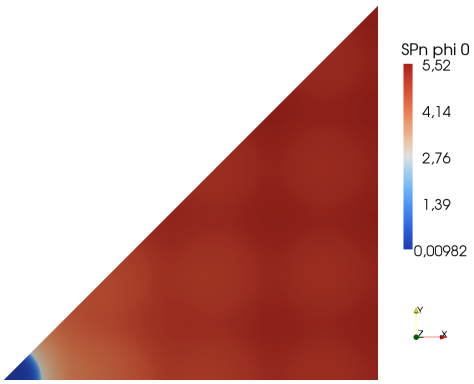


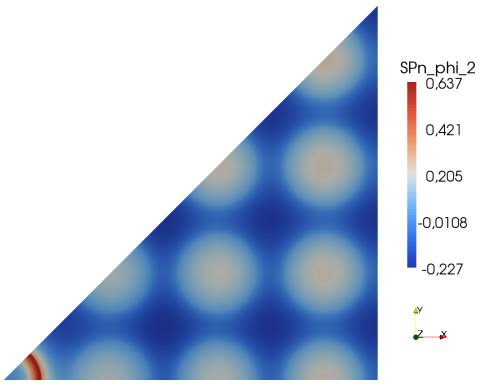
Figure 6.17: Solution of the Stankovski benchmark (SP_1).

These figures show the effort of higher-order SP_N approximations to recover the strong gradient of scalar flux at fuel pin boundaries via the SP_N fluxes ϕ^s (serving as correcting terms in (4.5)). The comparison of absorption rates

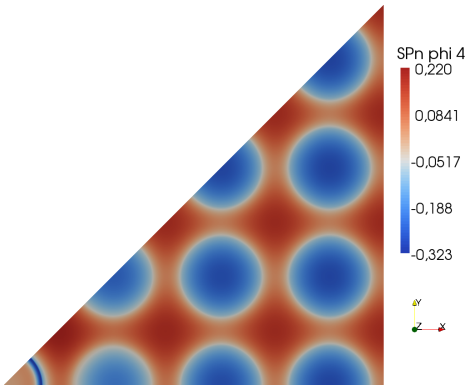
$$\int_{V_i} \sigma_{a,i}\phi(\mathbf{r})\,\mathrm{d}\mathbf{r}\,, \quad i = 1, 2, \ldots, 20,$$



(a) ϕ



(b) ϕ_2^s



(c) ϕ_4^s

Figure 6.18: Solution of the Stankovski benchmark (SP_5).

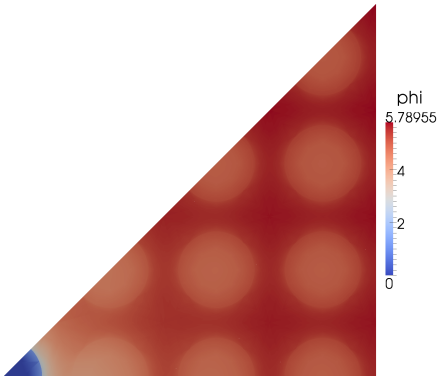


Figure 6.19: Solution of the Stankovski benchmark (S_8).

where V_i corresponds subsequently to `pin0`, `cell0`, `pin1`, `cell1`, etc., with the reference solution obtained from DRAGON (Fig. 6.20) indicates the limitation of the SP_N approximations when used outside their asymptotic range of theoretical validity (note that in some regions, the SP_5 errors are even slightly bigger than the SP_3 errors). It deserves mentioning that maximal difference between the reference region-wise integrated absorption rates and those obtained from the S_8 method implemented in Hermes2D was only 0.313% (more S_N examples in Hermes2D will be presented in Sec. 6.3.2).

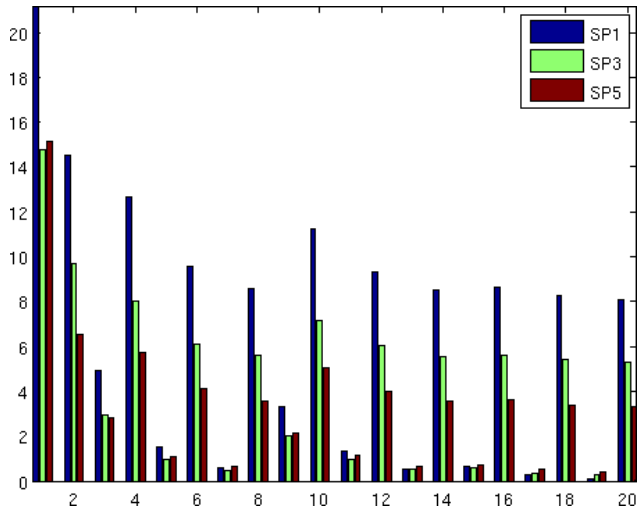


Figure 6.20: Comparison of various SP_N orders on cell-averaged absorption rate errors [%].

6.3.1.3 One-group hexagonal eigenvalue example

This example has been proposed in [63] to test the mixed finite element discretization of the SP_N equations. It represents a criticality eigenvalue problem

Region #	σ_t [cm^{-1}]	σ_{s0} [cm^{-1}]	σ_{s1} [cm^{-1}]	$\nu\sigma_f$ [cm^{-1}]
1	0.025	0.013	0.0	0.0155
2	0.025	0.024	0.006	0.0
3	0.075	0.0	0.0	0.0

Table 6.3: Material properties of the 1-group eigenvalue example.

with anisotropic scattering.

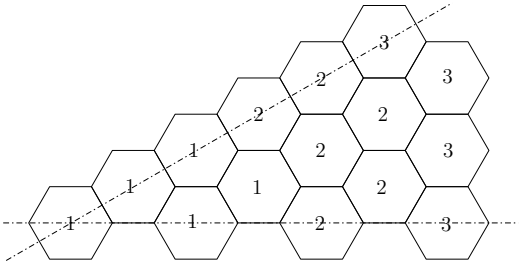


Figure 6.21: Geometry of the 1-group eigenvalue example. Each hexagonal assembly has side length 19 cm. Vacuum boundary conditions at the right boundary, reflective conditions along the dash-dotted lines.

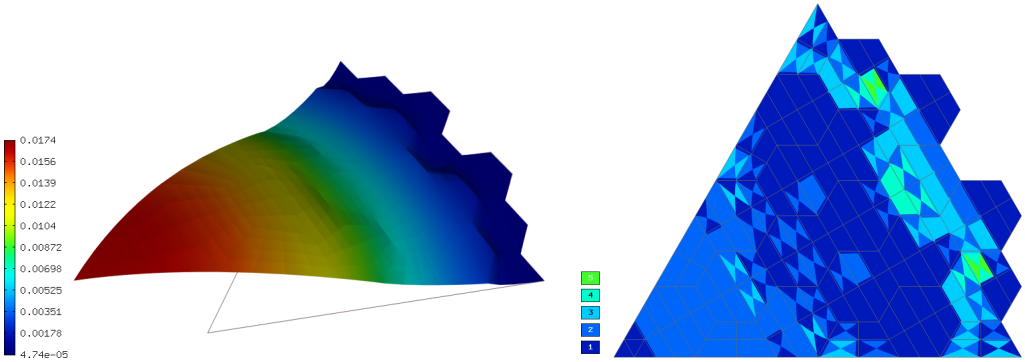


Figure 6.22: 1-group eigenvalue example – ϕ_0^s.

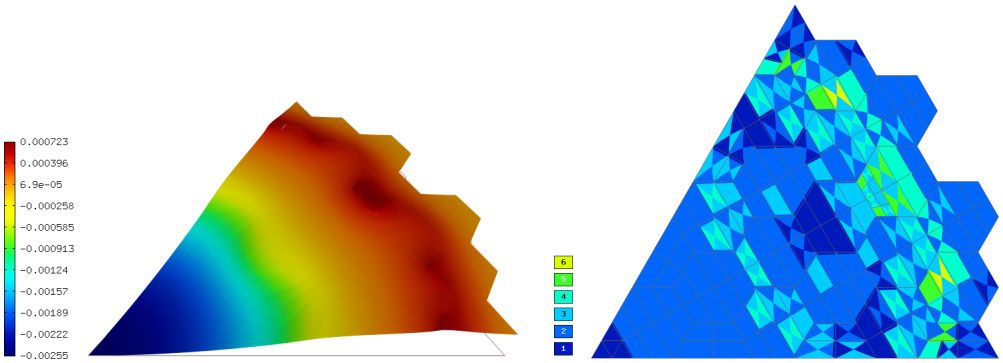


Figure 6.23: 1-group eigenvalue example – ϕ_2^s.

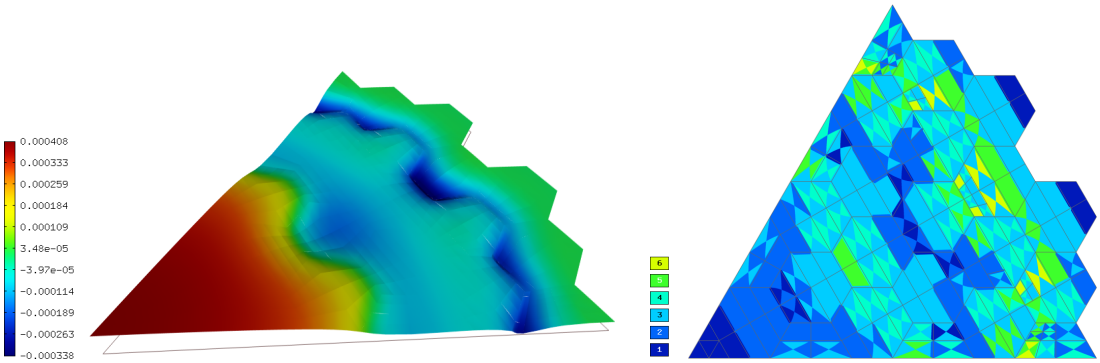


Figure 6.24: 1-group eigenvalue example – ϕ_4^s.

Figures 6.22 to 6.24 show the SP_5 solution (even order SP_5 fluxes) and the corresponding distribution of approximation orders and mesh sizes. At each adaptivity step, the Rayleigh-quotient iteration has been used to obtain the reference solution on the uniformly refined spaces, starting from orthogonal projection of the solution from previous adaptivity step onto these spaces (unit constant function has been used as initial shape of all moments). A decreasing tolerance has been used for the eigenvalue iteration, starting with 10^{-3} and decreasing all the way down to 10^{-10} by multiplication with 0.1 at each adaptivity step (the actual residual norm ratio to the initial one has been used to measure convergence:

$$\frac{(\|\mathbf{A} - \lambda^{(i)}\mathbf{B})\mathbf{x}^{(i)}\|}{(\|\mathbf{A} - \lambda^{(0)}\mathbf{B})\mathbf{x}^{(0)}\|},$$

where $\lambda^{(k)}$ is obtained from the Rayleigh quotient with $\mathbf{x}^{(k)}$).

The error in the final eigenvalue with respect to the reference solution from [63] was less than one milli-percent (pcm) as can be seen from the convergence curve in Fig. 6.25 (exactly 0.581 3 pcm) and the adaptivity convergence criterion $\|\mathrm{E}_{hp}^{\phi}\|_1 < 0.1\%$ was reached in 40.5 seconds. For comparison, using piecewise linear finite element approximation on the uniformly heavily refined mesh (Fig. 6.26) resulted in the eigenvalue difference of 2.784 70 pcm in 378.1 seconds (still using the direct solver UMFPACK with the same settings). Note that the reference solution in [63] was obtained using an approximation space with 118677 degrees of freedom, while the finest uniformly refined space needed to obtain the solutions depicted in the figures in this section using the *hp*-adaptivity in Hermes2D had 31898 degrees of freedom.

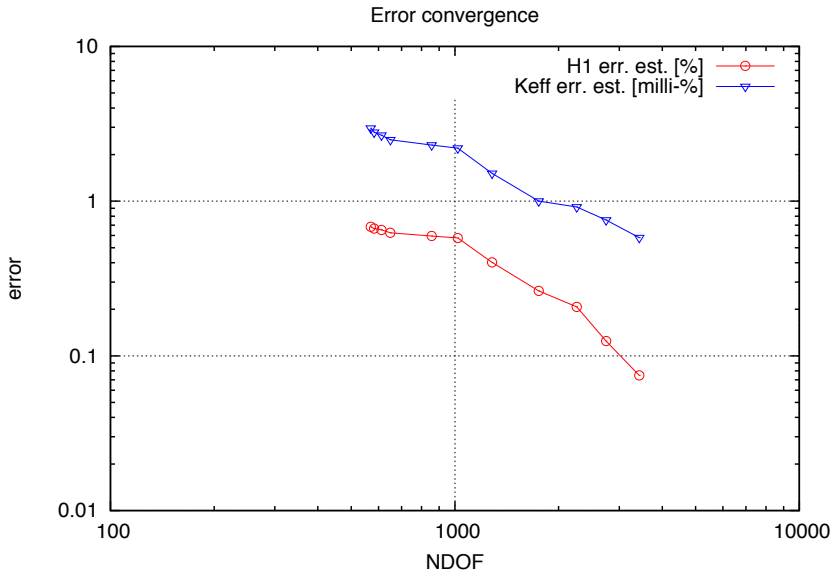


Figure 6.25: Adaptivity convergence curves for the 1-group eigenvalue example. The reference value $k_{\text{eff}} = 1.001271$ provided by [63].

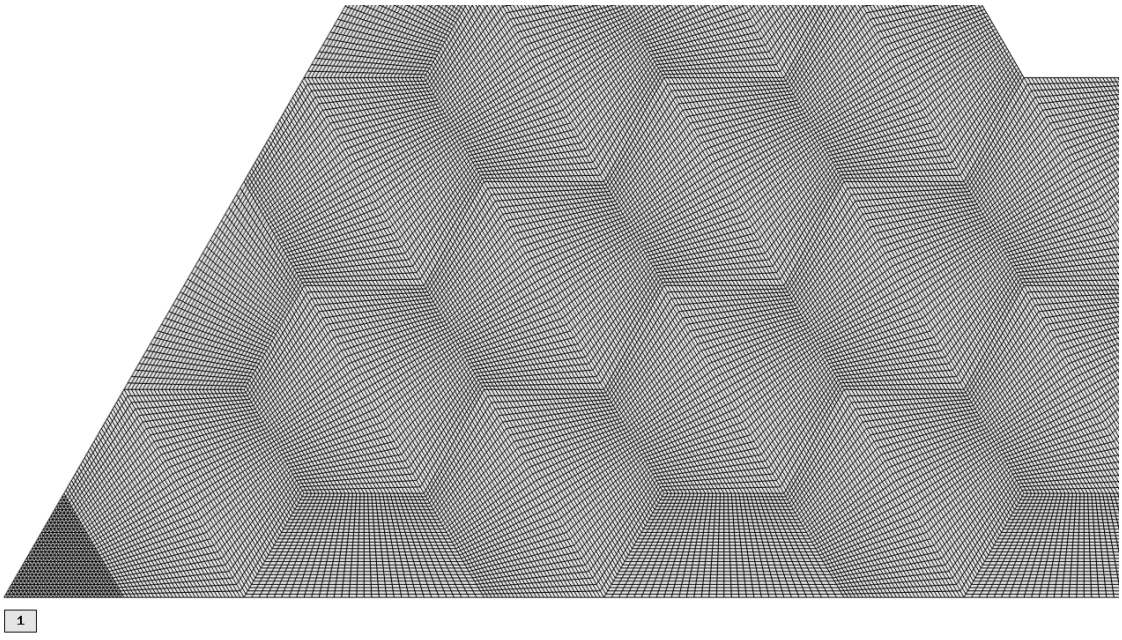


Figure 6.26: 1-group eigenvalue example (uniformly refined mesh).

6.3.1.4 Two-group WWER-440 criticality benchmark (diffusion)

This benchmark has been defined in [21] to test a nodal diffusion method against fine-mesh finite-difference solution of a two-group criticality eigenvalue problem for a core of the WWER-440 reactor with 1/12th reflectional symmetry. In the WWER-440 reactor, the control rods are represented by whole assemblies that push the regular fuel assemblies out of the core and replace their position. As a consequence, the multiplying medium with fission sources gets replaced by highly absorbing medium with no source, leading to strong solution gradients at the interface between the control rods and the adjacent assemblies. Also, high-magnitude solution gradients arise at the interface between the outer core assemblies and core reflector, which has been modelled in the benchmark by an additional shell of assemblies with special properties pertaining to the reflector. The non-smooth behavior of solution in these areas has been well captured by the *hp*-adaptivity procedure. The CPU time for this benchmark was 22.3 seconds.

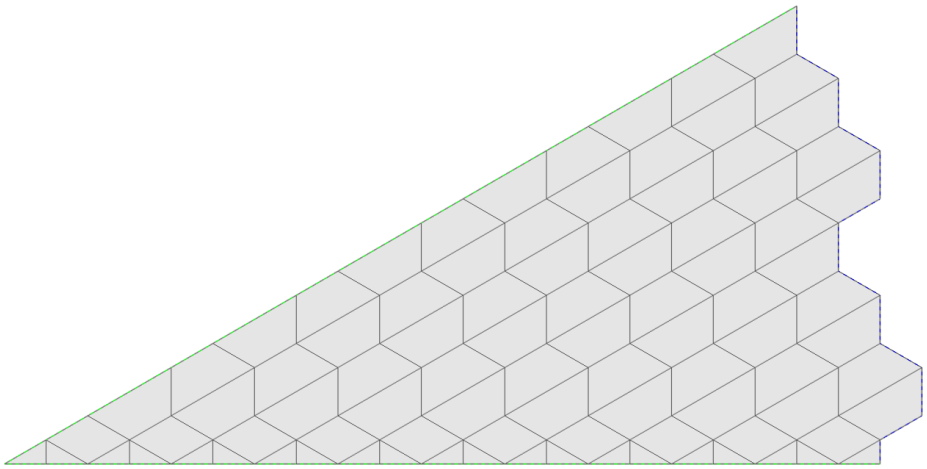
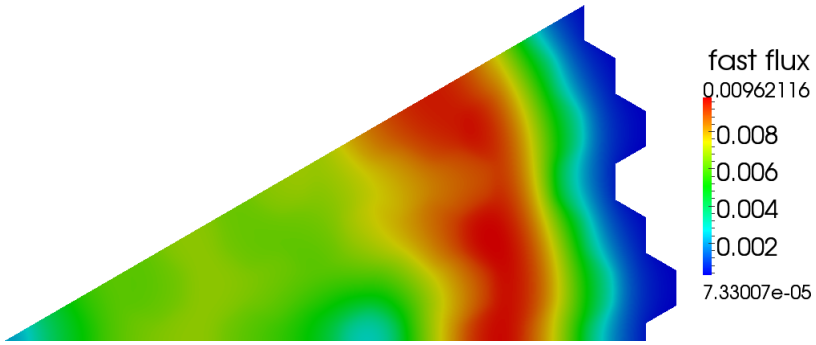
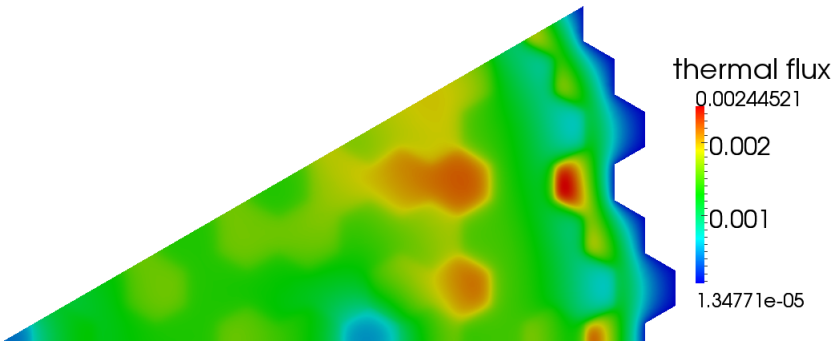


Figure 6.27: Initial mesh for the WWER-440 benchmark (reflective conditions on the diagonal and bottom line, vacuum conditions at the right boundary).

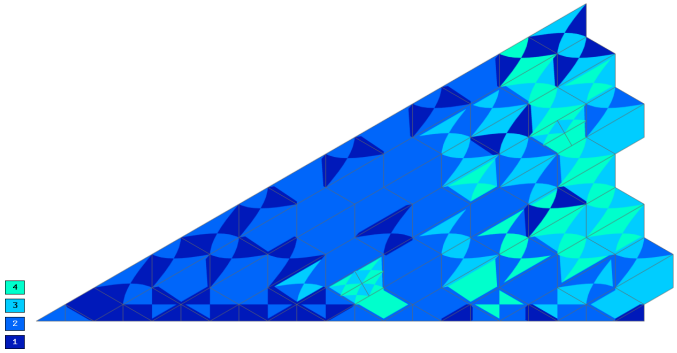


(a) ϕ^1

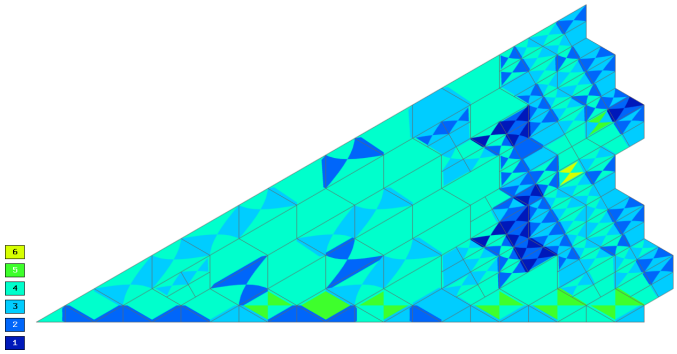


(b) ϕ^2

Figure 6.28: Solution of the WWER-440 benchmark (group scalar fluxes).



(a) $g = 1$



(b) $g = 2$

Figure 6.29: Approximation spaces for the WWER-440 benchmark.

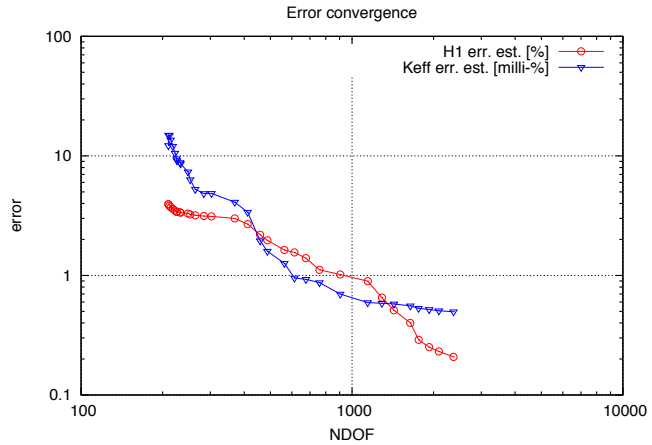


Figure 6.30: Adaptivity convergence curves for the WWER-440 benchmark. The reference value $k_{\text{eff}} = 1.00970$ provided by [21].

6.3.1.5 Four-group VHTR criticality benchmark (diffusion)

This benchmark of the criticality calculations in the azimuthally symmetric cylindrical geometry $(r - z)$ has already been implemented in Hermes2D as a standalone example when the author of this thesis joined the project. It served as

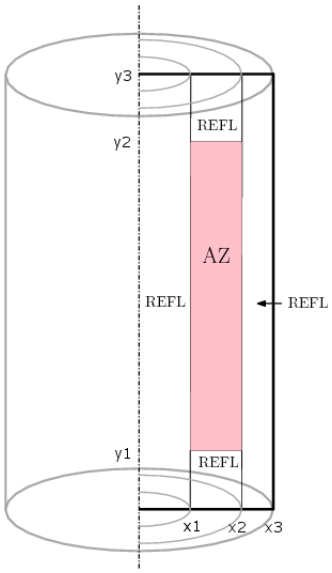
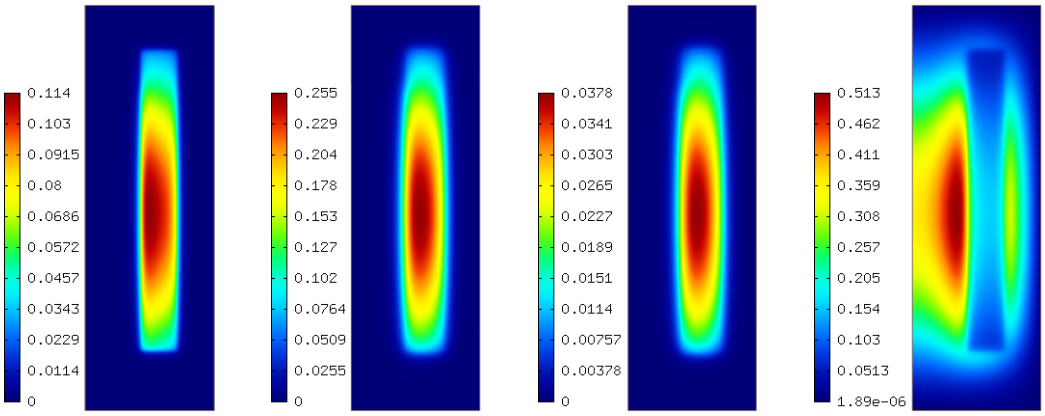


Figure 6.31: Geometry of the VHTR benchmark. Bold line encloses the rectangular region defining a computational domain for Hermes2D; vacuum boundary conditions are applied at these lines, while reflective condition is applied at the dash-dotted line. $x1 = 148\,\text{cm}$, $x2 = 242\,\text{cm}$, $x3 = 340\,\text{cm}$, $y1 = 158.5\,\text{cm}$, $y2 = 951.5\,\text{cm}$

one example of multimesh *hp*-adaptivity in [70]. Identifying $r = x$, $z = y$, the problem is conveniently placed into the Cartesian x-y system (gradient components stay the same, while all integrands need to be multiplied by $2\pi x$, i.e. the Jacobian determinant of the transformation between Cartesian and cylindrical coordinates). We remark that this holds also for integrals comprising the orthogonal projection and error calculation forms used during adaptivity (which was not taken into account in the original implementation, leading to less optimal convergence results).

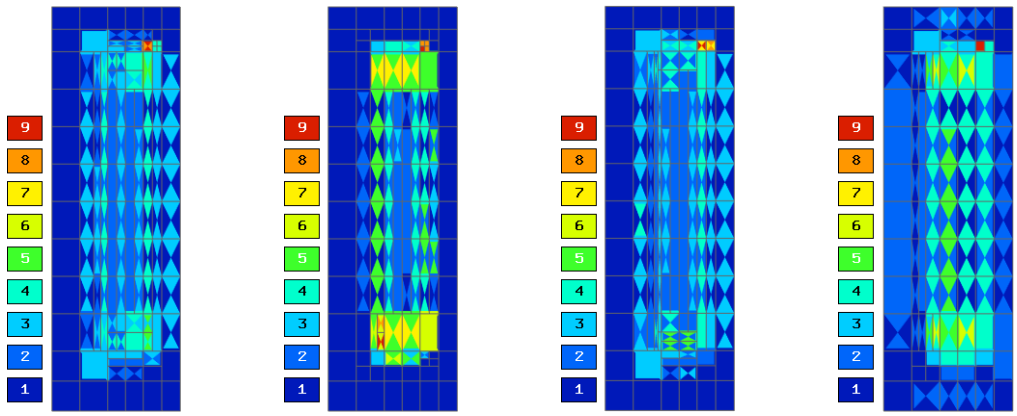
The results of the new version implemented in the neutronics framework and a more recent version of Hermes2D show the progress in the *hp*-adaptivity al-

gorithms and implementation – while 27903 degrees of freedom were required to reach $(H^1)^4$ error estimate of 0.0164% in [70], only 21010 degrees of freedom on the final uniformly refined space were required in the new version to reach comparable error estimate of 0.0172% (the corresponding coarse space solution had 5010 degrees of freedom as seen in Fig. 6.34).



(a) ϕ^1 (b) ϕ^2 (c) ϕ^3 (d) ϕ^4

Figure 6.32: Solution of the VHTR benchmark (group scalar fluxes).



(a) $g = 1$ (b) $g = 2$ (c) $g = 3$ (d) $g = 4$

Figure 6.33: Approximation spaces in the VHTR benchmark.

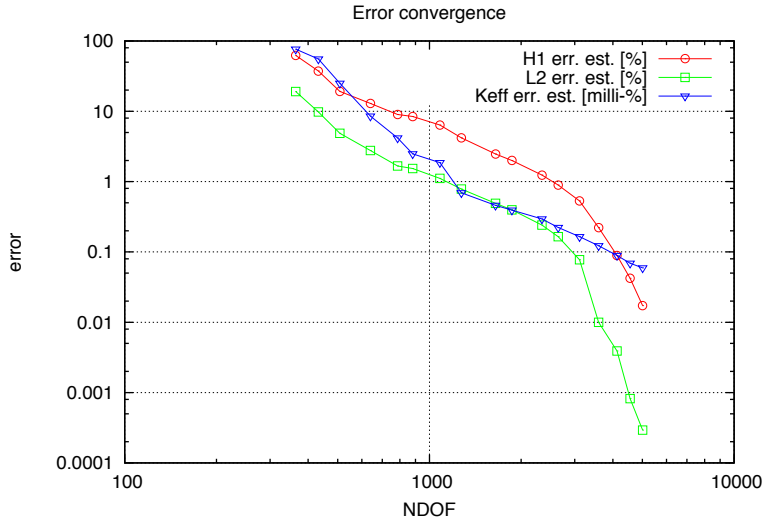


Figure 6.34: Adaptivity convergence curves for the VHTR benchmark. The reference k_{eff} value 1.1409144 was obtained by a reference calculation on a 3x uniformly refined mesh with uniform distribution of polynomial degrees (=4), with power method and convergence tolerance set to 5×10^{-11}. It slightly differs from the value 1.14077 reported in [70], where it is not clear, however, how the value was obtained.

6.3.2 S_N examples

6.3.2.1 Problem with exact solution

To verify the implementation of the S_N module, a simple solution

$$\psi(\mathbf{r}, \mathbf{\Omega}) = \psi(x, y, \Omega_x, \Omega_y) = \begin{cases} xy & \Omega_x > 0 \wedge \Omega_y > 0 \\ (1-x)y & \Omega_x < 0 \wedge \Omega_y > 0 \\ (1-x)(1-y) & \Omega_x < 0 \wedge \Omega_y < 0 \\ x(1-y) & \Omega_x > 0 \wedge \Omega_y < 0 \end{cases} \tag{6.6}$$

has been manufactured for the following two-dimensional transport problem with isotropic scattering, constant throughout a unit-square domain enclosed by vacuum:

$$\mathbf{\Omega} \cdot \nabla \psi(\mathbf{r}, \mathbf{\Omega}) + \sigma_t \psi(\mathbf{r}, \mathbf{\Omega}) - \frac{\sigma_s}{4\pi} \int_{\mathcal{S}_2} \psi(\mathbf{r}, \mathbf{\Omega}') \, \mathrm{d}\mathbf{\Omega}' = q(\mathbf{r}, \mathbf{\Omega}),$$

yielding the following source term:

$$q(\mathbf{r},\mathbf{\Omega}) = \Omega_x \begin{pmatrix} y-1 & \Omega_x<0 \wedge \Omega_y<0 \\ 1-y & \Omega_x>0 \wedge \Omega_y<0 \\ -y & \Omega_x<0 \wedge \Omega_y>0 \\ y & \Omega_x>0 \wedge \Omega_y>0 \end{pmatrix} + \Omega_y \begin{pmatrix} x-1 & \Omega_x<0 \wedge \Omega_y<0 \\ 1-x & \Omega_x<0 \wedge \Omega_y>0 \\ -x & \Omega_x>0 \wedge \Omega_y<0 \\ x & \Omega_x>0 \wedge \Omega_y>0 \end{pmatrix}$$
$$- \frac{\sigma_s}{4\pi} + \sigma_t \begin{pmatrix} xy & \Omega_x>0 \wedge \Omega_y>0 \\ (1-x)y & \Omega_x<0 \wedge \Omega_y>0 \\ (1-x)(1-y) & \Omega_x<0 \wedge \Omega_y<0 \\ x(1-y) & \Omega_x>0 \wedge \Omega_y<0 \end{pmatrix}.$$

It is easy to see that

$$\phi(\mathbf{r}) = \int_{\mathcal{S}_2} \psi(\mathbf{r},\mathbf{\Omega})\,\mathrm{d}\mathbf{\Omega} = \pi.$$

An exact solution $\psi(\mathbf{r},\mathbf{\Omega})$ with $\mathbf{\Omega} = \frac{\sqrt{2}}{2}[1,1]^T$ corresponding to the azimuthal angle $\varphi = 45°$ has been computed by Mathematica and is shown in Fig. 6.35.

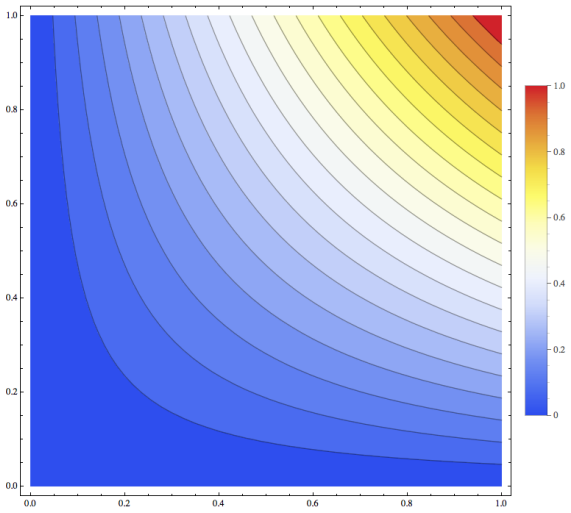


Figure 6.35: Manufactured solution problem – exact ψ_1.

Solution by Hermes2D for an arbitrary set of cross-sections $\sigma_t = 1\,\mathrm{cm}^{-1}$, $\sigma_s = 0.5\,\mathrm{cm}^{-1}$ is presented in figures 6.36 (scalar flux) and 6.36 (angular fluxes in first four directions of the S_4 set; same set of solutions have been obtained in the remaining directions, as expected from (6.6)).

Note that although the solution does not depend on the choice of σ_t and σ_s, $c = \sigma_s/\sigma_t$ will influence the convergence rate of the source iteration process implemented in Hermes2D, as noted in Sec. 3.4.3. Figure 6.36 shows the solution of

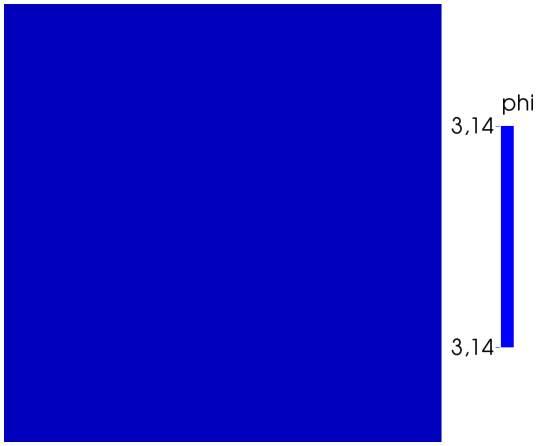
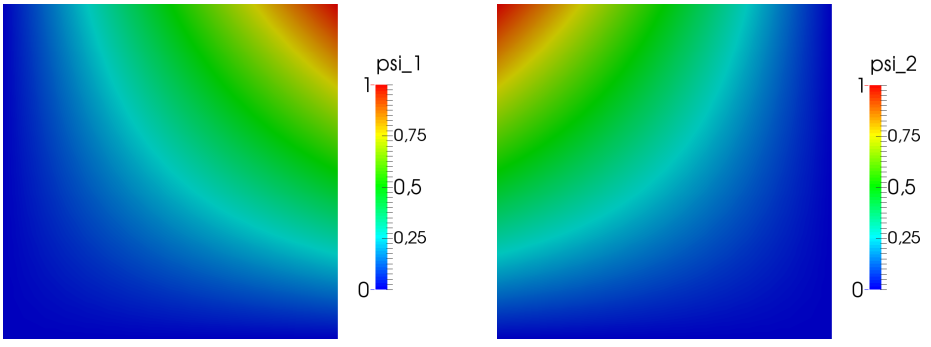
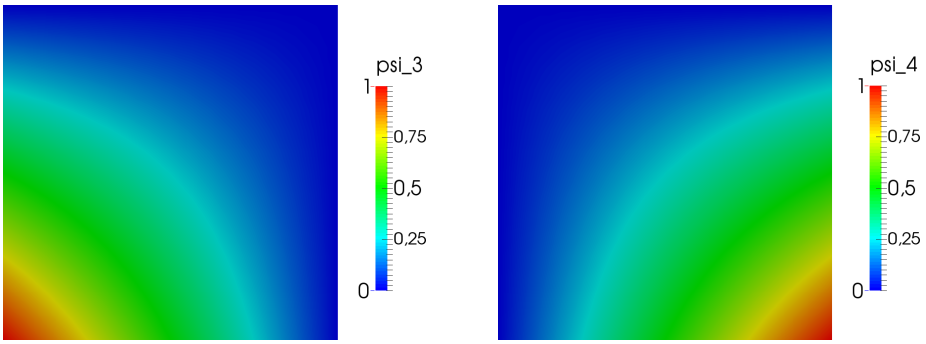


Figure 6.36: Manufactured solution problem – converged scalar flux.



(a) ψ_1 (b) ψ_2



(c) ψ_3 (d) ψ_4

Figure 6.36: Manufactured solution problem – converged angular fluxes in first direction within each quadrant of $\mathcal{S}_2$.

source iteration converged with tolerance 10^{-10} (19 iterations), while the solution converged with tolerance 10^{-5} (9 iterations) is shown in Fig. 6.37. As indicated by Theorem 5, the convergence rate depends on the product $c\sigma_t = \sigma_s$; 17 iterations were needed to converge the solution for $\sigma_t = 100\,\mathrm{cm}^{-1}$, $\sigma_s = 50\,\mathrm{cm}^{-1}$ with the tolerance set to the same value 10^{-5} (Figure 6.38).

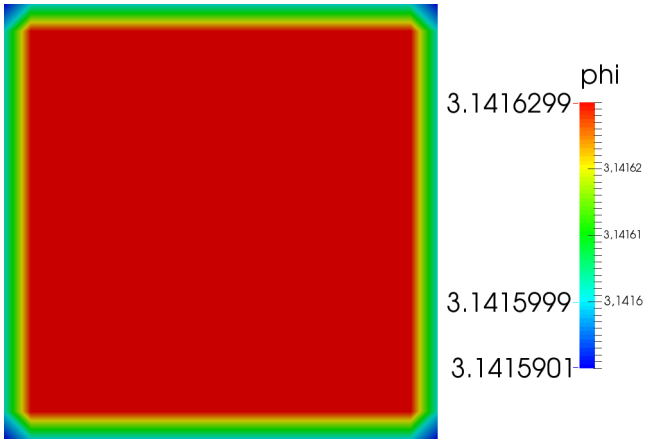


Figure 6.37: Manufactured solution problem – scalar flux from angular fluxes converged to within tolerance 10^{-5}.

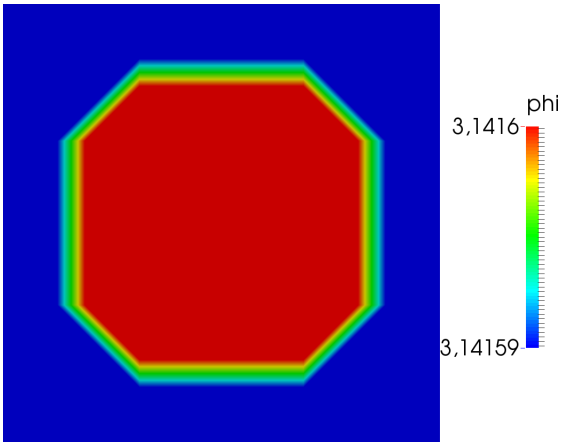


Figure 6.38: Manufactured solution problem – scalar flux from angular fluxes converged to within tolerance 10^{-5} (different cross-section set).

6.3.2.2 Watanabe-Maynard Problem 1

This benchmark has been presented in [120] to demonstrate ray-effect mitigation capabilities of special methods developed in the paper. A uniform isotropic source is placed at the center of a vacuum-surrounded domain and separated from the scattering part of the domain by a layer of vacuum as illustrated in Fig. 6.39.

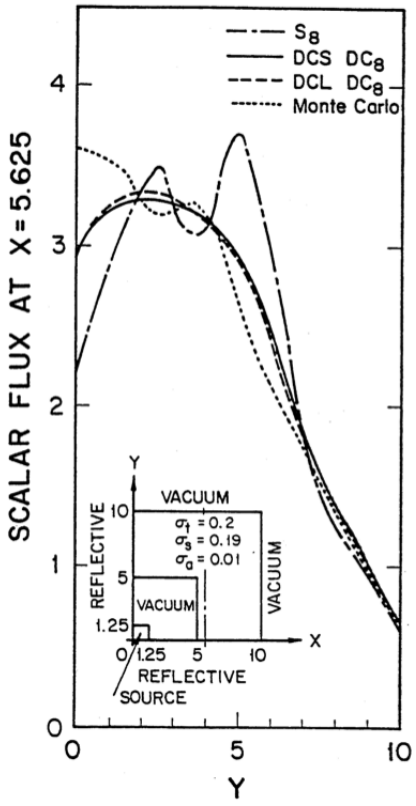


Figure 6.39: Geometry and reference solutions of the Watanabe-Maynard problem (from [120]). Source strength is $6.4\,\mathrm{cm}^{-3}$·sec.

While the ray-effects are clearly present in our S_8 solution (figures 6.41, 6.42), which has not been treated to diminish these oscillations in any special way, these appear to be less pronounced than when using the standard S_8 ordinates set (line "S8" in Fig. 6.39).

Note also that ray effects become more visible when refining the mesh (Fig. 6.42), confirming the importance of keeping the mesh refinement in harmony with the

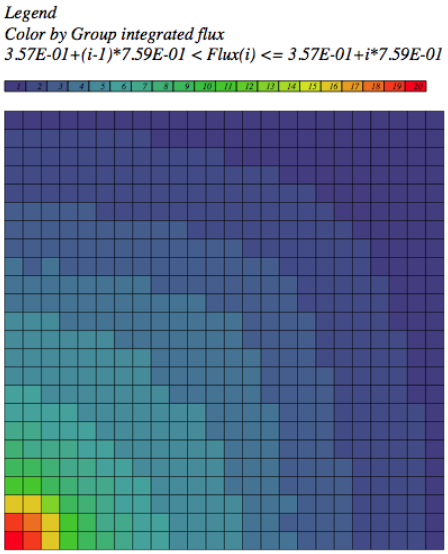
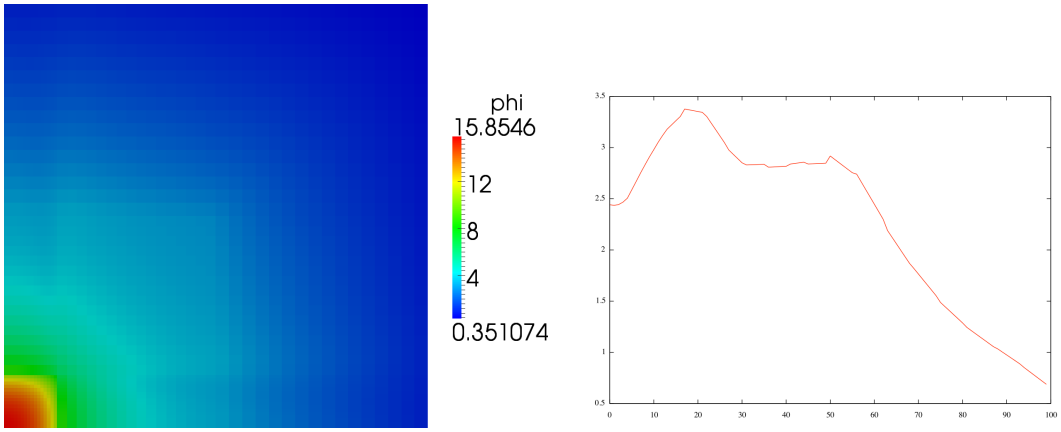


Figure 6.40: Solution of the Watanabe-Maynard problem by DRAGON. Values span the range $[3.57 \times 10^{-1}, 15.537]$.

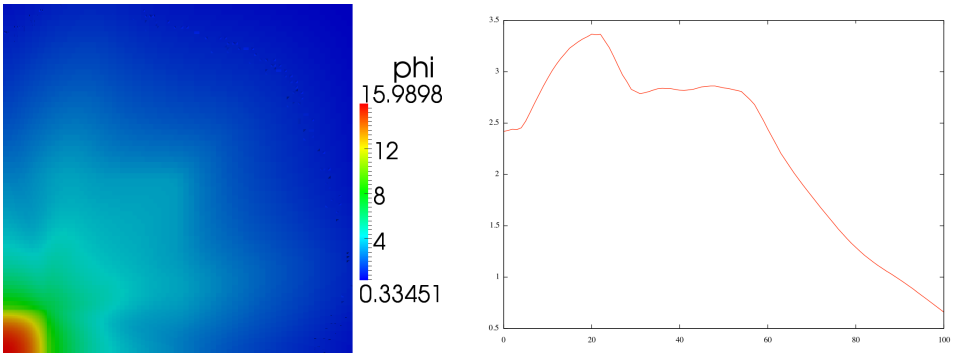


(a) scalar flux

(b) scalar flux at $x = 5.625\,\mathrm{cm}$

Figure 6.41: Solution of the Watanabe-Maynard problem, using the mesh from Fig. 6.40 with one level of uniform refinement and $DG(0)$ elements.

increase of the number of directions. For comparison, Fig. 6.40 shows the solution provided by the code DRAGON using the method of characteristics using the setup from [80, Sec. 6.4.3].



(a) scalar flux (b) scalar flux at $x = 5.625\,\mathrm{cm}$

Figure 6.42: Solution of the Watanabe-Maynard problem, using the mesh from Fig. 6.40 with two levels of uniform refinements and $DG(0)$ elements.

6.4 Coupled code system for quasi-static whole-core calculations

A fully three-dimensional multigroup neutron diffusion solver was needed for the purposes of the project "Project TA01020352 – Increasing utilization of nuclear fuel through optimization of an inner fuel cycle and calculation of neutron-physics characteristics of nuclear reactor cores", which the author of this thesis participated in during his doctoral studies. As the 3D version of Hermes has not been released yet, another finite-element library has been chosen and the experience obtained from developing the neutronics solvers within Hermes2D was utilized to develop a similar framework using this library as a backend. Namely, the FE matrix assembly is handled by the well-established open-source system FEniCS [77, 78] and its linear algebra backend PETSc [10]. PETSc library and its fork focusing on solving large sparse eigenvalue problems, SLEPc [57], is also primarily used for solving the assembled algebraic problems (mainly the generalized eigenvalue problems of form (3.5) that are of most importance in this project; cf. also the introduction to Sec. 2.2.6). This has the advantage that parallel assembly and solution using MPI is almost automatic (provided an appropriate solver/preconditioner is being used).

The code operates along the scheme shown in Fig. 6.43 and is written in

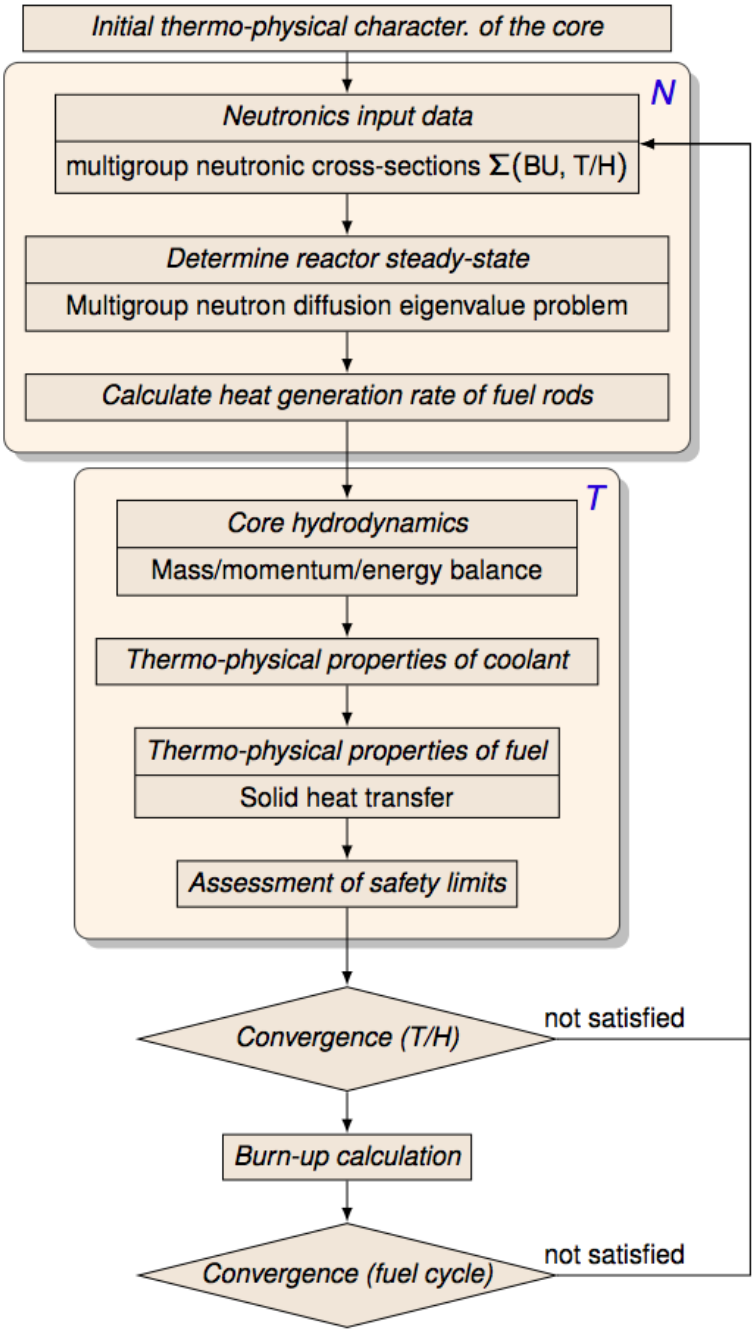


Figure 6.43: Coupled code run scheme.

Python 2.7 with critical parts (typically where loops over mesh cells are needed) in C++ as SWIG extension modules. The ability to mix Python with C++ extensions as well as availability of Python modules that can achieve C++-like performance if used in appropriate way (Numpy for operations with large data arrays, mpi4py for MPI-related tasks) make the code sufficiently fast and flexible at the same time. Thorough description of the code and its results on standard industry benchmarks formed the content of three research reports that were accepted by the project issuing agency (one final report is currently in preparation). To keep the scope of this thesis reasonable, just a sample of these results for one such benchmark – the steady-state part of the OECD/NEA MOX-UO2 benchmark [69] – is presented below.

The problem was solved using the Jacobi-Davidson generalized eigensolver [98] with a PETSc implementation of BiCGStab(ℓ) [106] as an inner solver (using the default $\ell = 2$). It has been demonstrated already in [106] (and analyzed in many later papers, see e.g. [58] and references therein) that the Jacobi-Davidson method is remarkably robust with respect to accuracy of the solution of the inner solution phase. Therefore, by setting a fixed number of inner iterations to 20 and employing a smoothed aggregation algebraic multigrid preconditioner with 1 pre- and 1 post-smoothing Richardson iterations (from the Trilinos ML 6.0 package wrapped available through PETSc interface), eigenvalue convergence within at most 10 outer iterations was achieved in every feedback step. 34 feedback iterations were required in the hot full power calculation to converge all fields of interest to within 10^{-4} relative difference from previous iteration, leading to a total calculation time of 1053 seconds. The calculation ran on a server with 12 Intel Xeon E5645 / 2.40 GHz processors, 12 MB of L3 cache, 24 GB RAM and Debian 3.2.51-1 64bit operating system.

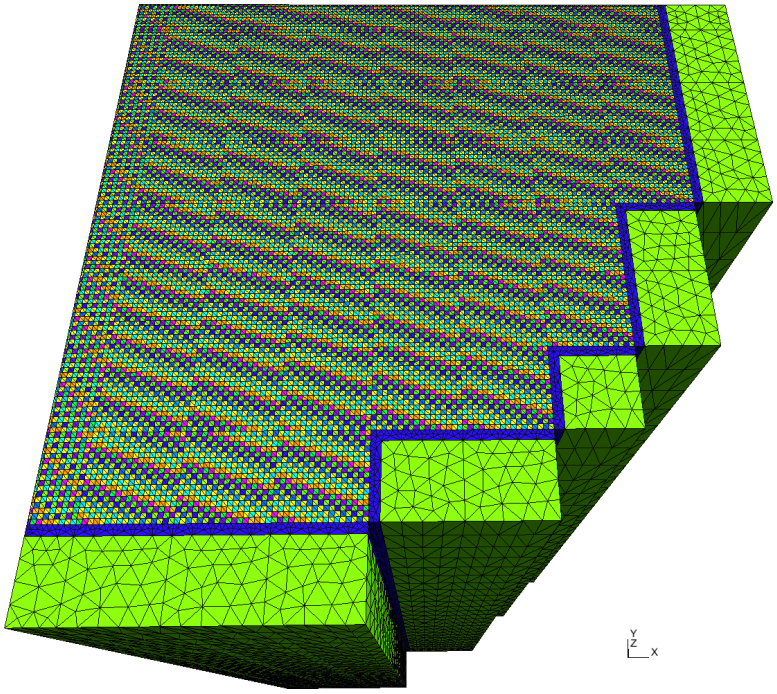


Figure 6.44: Mesh for the OECD/NEA MOX-UO2 benchmark (generated by GMSH [53]). Colored by regions that can be assigned different properties (i.e., fuel pins in the core interior; note that assembly-homogenized cross-sections are given in the benchmark, but T/H feedbacks are actually calculated pin-by-pin).

	%PWE	c_b	T_D	ρ_m	T_m	ρ_m^o	T_m^o
CORETRAN	0.31	1647	908.4	706.1	581.0	658.5	598.6
CORETRAN 4/FA	0.26	1645	908.4	706.1	581.0	658.5	598.6
EPISODE	0.40	1661	846.5	701.8	582.6	697.4	585.5
NUREC	0.31	1683	827.8	706.1	581.1	661.5	598.7
PARCS2G	ref	1679	836.0	706.1	581.3	662.1	598.8
SKETCH-INS	1.04	1675	836.6	705.5	580.9	659.6	598.9
Results	1.35	1699	839.0	702.1	582.3	660.1	599.1

Table 6.4: OECD/NEA MOX-UO2 benchmark (hot full power case) – comparison of results with various nodal methods from [69].
$PWE = \frac{\sum_{a\in\text{assemblies}}|P_a - P_a^{\text{ref}}|P_a^{\text{ref}}}{\sum_{a\in\text{assemblies}} P_a^{\text{ref}}}$, $P_a \ldots$ assembly integrated power
c_b: critical boron concentration, T_D: Doppler (fuel) temperature, ρ_m: moderator density, T_m: moderator temperature, $\cdot^o$: outlet average value.

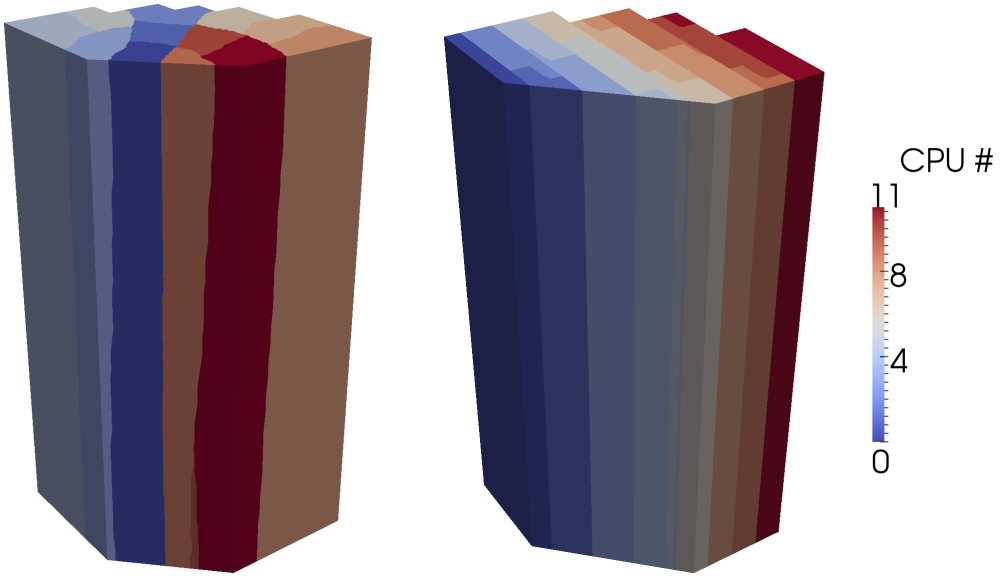


Figure 6.45: Mesh distribution on 12 CPUs – automatic distribution by the SCOTCH library (used by FEniCS for assembling the FE matrices) on left, manual redistribution for efficient parallel T/H calculations on right (i.e., complete axial channels from core inlet to outlet are available at each processor).

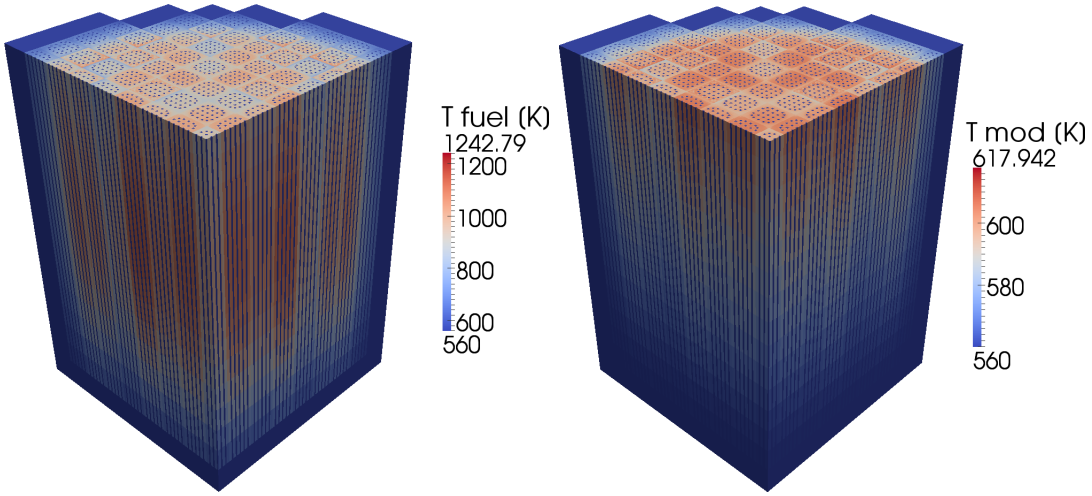
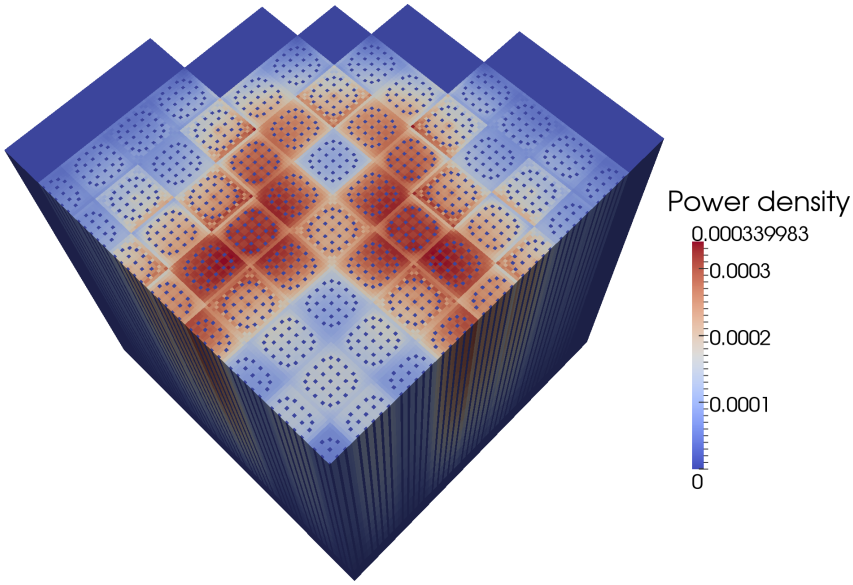
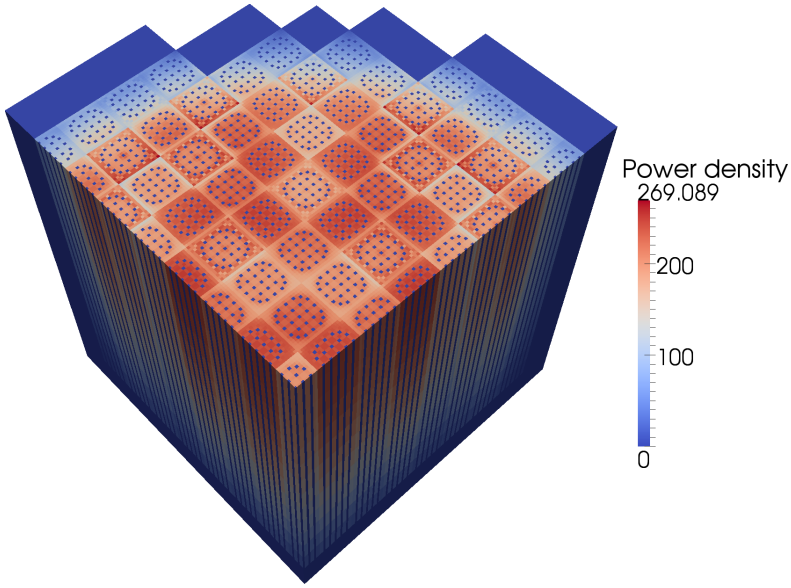


Figure 6.46: T/H fields distribution.

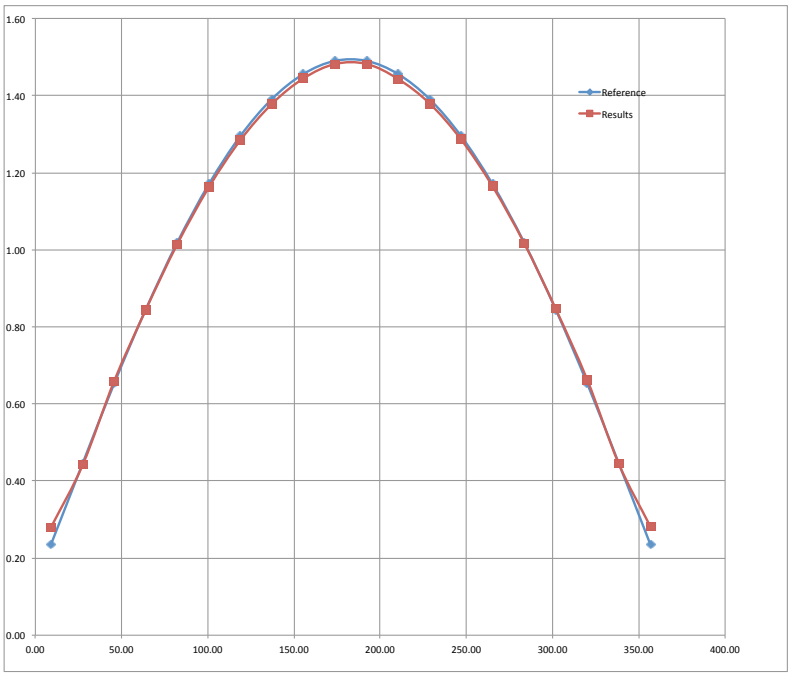


(a) hot zero power case

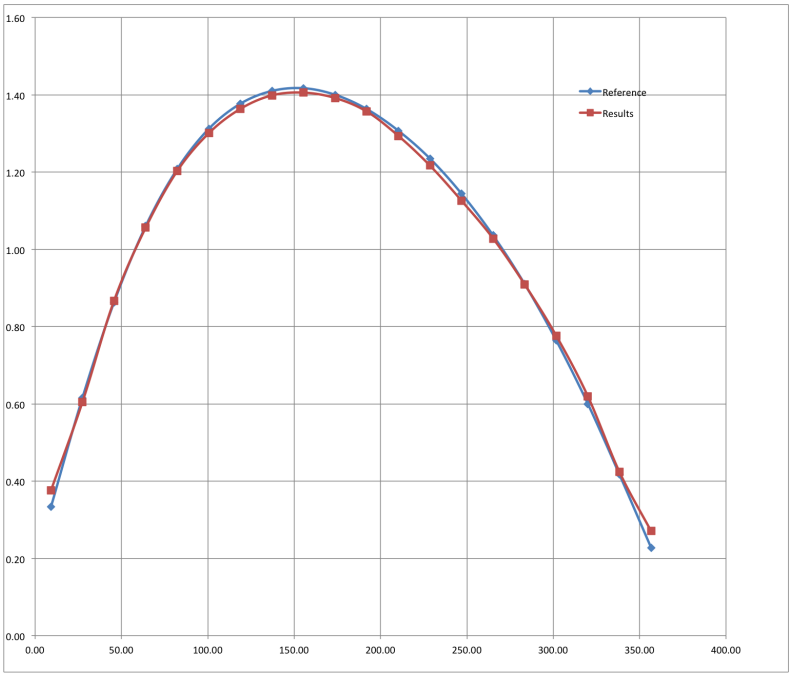


(b) hot full power case

Figure 6.46: Core-wide power distribution.



(c) hot zero power case



(d) hot full power case

Figure 6.47: Axial power distribution.

7

Summary

The overarching subject of this work was computational modeling of neutron transport relevant to development and analysis of computer codes for simulating the long-term operation of nuclear facilities. Its focus was on deterministic models originating in the linear Boltzmann's transport equation – in a general three-dimensional setting, an integro-differential equation describing the evolution of neutron distribution as a function of time, three spatial, two angular and one energy variable. In particular, the steady state form of the equation was studied in this work, forming the basis for both steady and transient neutron transport calculations. The general mathematical model was presented in Chap. 2, together with a review of known results about the solvability and well-posedness of two main problems arising in neutron transport – the fixed source problem and the reactor criticality (eigenvalue) problem.

High-dimensionality and complicated structure of the equation for real-world problems preclude in most situations an analytical solution and present serious difficulties when a computational solution is attempted. Several dimension-reduction techniques thus need to be employed along with appropriate numerical solution schemes. These techniques were reviewed in Chap. 3, starting with the classical discretization of energetic dependence – the multigroup method.

Two most widely used methods for approximating angular dependence (the S_N and P_N methods) were presented next. We demonstrated how these two methods – well known for a long time, but typically presented independently of each other – can be described using a single abstract mathematical framework. Sec. 3.3.1

showed the standard derivation of the P_N equations as an orthogonal projection of the continuous neutron transport equation onto a subspace of $L^2(\mathcal{S}_2)$ spanned by suitable functions of angular variable (the spherical harmonics). Sec. 3.4.2 showed how the S_N equations can also be obtained by projecting the NTE onto a suitable subspace of $L^2(\mathcal{S}_2)$, at least in the case of isotropic scattering. This time, however, the projection is no longer orthogonal. Viewing the S_N approximation as a subspace projection makes the properties of the continuous NTE directly translate to the approximate method. As a particular example, Sec. 3.4.3 outlined how this fact could be used to prove convergence of the basic iterative method for actual solution of the S_N equations – the source iteration.

The part devoted to angular approximation methods was concluded by introducing a concrete angular quadrature set for the S_N approximation, satisfying various requirements uncovered during the description of the approximation. This quadrature set was later used in an actual computer code for solving the S_N equations, described in Chap. 6. In order to obtain this final neutron transport solver, the spatial dependence in the NTE had to be discretized. For this purpose, a general discontinuous Galerkin method was applied to the S_N equations in Sec. 3.5.2 to obtain the weak formulation in a form suitable for implementing into common finite element libraries.

The attention was then turned to the P_N approximation. The diffusion equation, a standard tool for whole-core nuclear reactor calculations that can be obtained from the lowest order P_1 approximation under well-understood assumptions, has been recalled in Sec. 3.3.5 and its weak form suitable for finite element spatial discretization in Sec. 3.5.3. Unlike the S_N approximation, standard continuous Galerkin finite elements can be used for discretizing the diffusion equation and a symmetric and positive definite system of algebraic equations is obtained by the procedure in the monoenergetic case. This allows to use matured methods for solving such algebraic systems, like the conjugate gradient method preconditioned by algebraic multigrid.

Some recent nuclear reactor problems involving higher amount of material and geometric heterogeneity, where diffusion no longer provides an adequate approximation and using S_N approximation with sufficient angular resolution would be prohibitive in terms of computer resources, have increased the interest of nuclear

engineers in the simplified P_N approximations – higher order approximations of neutron transport that preserve the simple and numerically convenient structure of the diffusion equation. In Chap. 4, we repeated the classical derivation of these equations by a straightforward extension of the approach used to derive the diffusion equation from the P_1 equations to higher orders. However, this approach is rigorous only in 1D case where the basis for the P_N method is represented by the Legendre polynomials. An ad-hoc generalization is needed in multiple dimensions as described in Chap. 4, although, as also recalled in this chapter, asymptotic and variational analyses that have been performed over the years eventually provided more sound mathematical grounds for this derivation. One of the contributions of this thesis to the vast body of existing SP_N knowledge is the strengthening of theoretical foundations also for the numerical methods for solving the SP_N equations – Chap. 4 was concluded by formulating the weak form of the SP_N system and proving its well-posedness for $N = 3, 5, 7$ (the orders that generally have the most impact on accuracy while keeping the computational requirements on acceptable levels) with an obvious extendability to higher orders.

In Chapter 5, we were concerned with the derivation of the SP_N equations itself. An alternative set of equations corresponding to the P_N approximation was introduced, using instead of the tesseral spherical harmonic basis as in the P_N approximation a basis formed by special linear combinations of these functions as a starting point. These Maxwell-Cartesian surface spherical harmonics form components of totally symmetric and traceless Cartesian tensors and provide a direct generalization of Legendre polynomials to multiple dimensions. The resulting system of MCP_N equations – first order partial differential equations for the unknown moment tensors of increasing rank – was derived and explicitly written for the case of $N = 3$. It was then demonstrated that in the interior of a homogeneous region, operations analogous to those used to derive the diffusion equation from the P_1 approximation and tensor structure of the equations can be used to manipulate the MCP_3 system into a set of second-order partial differential equations. Moreover, it was shown that the zero-th order solution moment (scalar flux) satisfies a weakly coupled set of two diffusion-like equations. Chap. 5 was concluded by exhibiting the equivalence of this set with the SP_3 equations presented in Chap. 4.

The final chapter addressed actual implementation of the neutron transport approximations described in previous chapters as neutronics modules in existing finite element frameworks. We started by the S_N and adaptive multigroup SP_N modules in the Hermes2D library. It was explained why Hermes2D is a particularly suitable tool for developing adaptive solvers of coupled systems of PDE's (in 2D), i.e. its advanced approach to *hp*-adaptivity and the multimesh assembling capability, which allows to approximate each solution component using its own adapted approximation space and mesh without introducing any interpolation errors. The basic idea behind multimesh assembling was used by the author and his colleague L. Korous in the implementation of discontinuous Galerkin assembling procedure, described in Sec. 6.1.2. Regarding the neutronics modules, this newly added feature was utilized when writing the S_N module. We also presented a simple modification of standard element error indicator used by Hermes2D to drive the *hp*-adaptivity that is more suitable for quantifying error in scalar flux (the physical quantity of interest) rather than in the artificial SP_N fluxes.

Actual examples of solution of typical neutron transport problems by the developed code constituted the second part of the final chapter. Both fixed source and eigenvalue problems were solved using diffusion and higher order SP_N approximations and limitations of the various models were exhibited. Errors arising from spatial approximation were minimized by employing the *hp*-adaptivity of Hermes2D. Two problems verifying the implementation of the S_N model were also presented in this part. Even though the Hermes library supports only two-dimensional problems, the experience gained by implementing neutronics modules into it was utilized in the development of a 3D multigroup diffusion solver with nonlinear thermal/hydraulic feedback coupling for the purposes of a major research project investigated at author's department. Sample results of this solver concluded the thesis.

Directions for future research

There are several directions in which the topics of this thesis may be further explored. In Chap. 3, we showed how the S_N approximation can be formulated as a projection onto a Hilbert subspace of $L^2(\mathcal{S}_2)$. The derivation was limited to

the case of isotropic scattering, so that a proper basis for the projection subspace could be established. The study of the possibility of using Dirac delta distributions instead of the piecewise constant functions ι_m could lead to extension of the analysis to general anisotropic scattering, as remarked at the end of Sec. 3.4.2.1. The Hilbert space projection approach certainly opens more ways of analyzing the behavior of numerical methods by the means of functional analytic tools than the simple convergence analysis of the S_N source iteration presented in Sec. 3.4.3 – see for instance [67] for an application to abstract preconditioning by Riesz mappings.

In Chap. 5, we derived the general form of the MCP_N equations, but then proceeded to a coupled diffusion-like system of equations for scalar flux only for the MCP_3 equations and isotropic scattering. The possibility of extension to higher orders and anisotropic scattering is apparent from the procedure, but would be more technically involved. However, using computer algebra systems like Mathematica could greatly simplify and automate this task (as it did for the SP_N derivation, cf. Appendix C). The search for interface and boundary conditions remains the largest open question regarding the coupled diffusion-like system derivable from the MCP_N approximation (see Sec. 5.4.3). Note that the approach used in [105] is based on the classical general Gauss-Ostrogradski pill-box argument and requires to express the second-order moment tensor $\psi^{(2)}$ in terms of ϕ (which constitutes another puzzling equation, eq. (7), in the paper). In a preliminary investigation, we were able to derive [105, Eq. (7)] as a scalar part of the scalar-vector-tensor decomposition of $\psi^{(2)}$ ([26]), which is the complete representation of the tensor under the assumption of $\psi^{(1)}$ and $\psi^{(3)}$ being curl-free (conservative). This indicates one direction that may be taken in the quest for complete understanding of Selengut's paper (and formulating a practical way of obtaining P_N scalar flux from solving much smaller system of elliptic equations).

A large space for further work lies in the area of adaptive solution methods for the NTE. While *h*-adaptivity for the S_N equations has received attention in several recent papers (see e.g. [41, 45, 95] for adaptivity based on a-posteriori estimation of global L^2 norm of solution error or [74, 119] for a method based on goal oriented adaptivity), [46] appears to be the first paper aimed at employing *hp*-adaptivity for the S_N equations. Prevalence of h-adaptivity and use of $p = 1$

(linear) finite element spaces is caused partly by the difficulty of implementation of an *hp*-adaptive FE code itself, partly by the fear of the well known limited regularity of the exact solution of the NTE even for smooth input data. However, similarly to the experience with hp-adaptive methods in different fields, the limitation of asymptotic convergence rate (as $h \to 0$ where h is the diameter of the largest element in the mesh) dictated by a-priori error estimates involving solution regularity typically doesn't appear until very late in the mesh refinement process or at all ([118]). Hence, utilizing higher order approximations still makes sense to accelerate pre-asymptotic convergence rate as much as possible. Implementation of an S_N solver in the general hp-FEM framework provided by the Hermes2D library (see Chap. 6) could be used as a basic building block for future investigations in this direction. Note that some kind of angular adaptivity should also be considered in conjunction with spatial adaptivity for the S_N equations in order to keep the ray-effects under control (cf. Sec. 6.3.2.2).

Appendix A

Spherical harmonics

The (tesseral) spherical harmonic function of degree n and order m ($n \in \mathbb{N}_0,\ m \in \mathbb{Z}, 0 \leq |m| \leq n$) is defined as ([49, Sec. 3.13])

$$Y_n^m(\mathbf{\Omega}) = Y_n^m(\vartheta, \varphi) = \sqrt{(2 - \delta_{m0}) \frac{2n+1}{4\pi} \frac{(n-|m|)!}{(n+|m|)!}} P_n^{|m|}(\cos\vartheta) T_m(\varphi), \quad \text{(A.1)}$$

or, if we denote $\mu = \cos\vartheta$ the cosine of the polar component of the direction vector $\mathbf{\Omega}$ (see Fig. A.1),

$$\sqrt{(2 - \delta_{m0}) \frac{2n+1}{4\pi} \frac{(n-|m|)!}{(n+|m|)!}} P_n^{|m|}(\mu) T_m(\varphi);$$

$\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ otherwise is the Kronecker delta symbol,

$$P_n^m(\mu) = \sqrt{(1-\mu^2)^m}\ \frac{d^m P_n(\mu)}{d\mu^m}, \quad m > 0 \quad \text{(A.2)}$$

are the *associated Legendre functions*,

$$T_m(\varphi) = \begin{cases} \cos m\varphi & m \geq 0, \\ \sin|m|\,\varphi & m < 0 \end{cases}$$

and $P_n(\mu) = P_n^0(\mu)$ is the n-th member of the system of Legendre polynomials:

$$\begin{gathered} P_0(\mu) = 1, \quad P_1(\mu) = \mu, \quad P_2(\mu) = \frac{1}{2}(3\mu^2 - 1), \quad P_3(\mu) = \frac{1}{2}(5\mu^3 - 3\mu) \\ (2n+1)\mu P_n(\mu) = (n+1)P_{n+1}(\mu) + nP_{n-1}(\mu), \quad n = 1, 2, \ldots \end{gathered} \quad \text{(A.3)}$$

Spherical harmonics form a complete orthonormal system on $L^2(\mathcal{S}_2)$ with respect to the standard inner product

$$(\psi, \varphi)_{L^2(\mathcal{S}_2)} = \int_{\mathcal{S}_2} \psi(\mathbf{\Omega})\varphi(\mathbf{\Omega})\,\mathrm{d}\mathbf{\Omega}\,,$$

that is

$$\begin{aligned}\int_{\mathcal{S}_2} Y_n^m(\mathbf{\Omega})Y_n^m(\mathbf{\Omega})\,\mathrm{d}\mathbf{\Omega} &= \int_0^{2\pi}\mathrm{d}\varphi \int_0^{\pi}\sin\vartheta\mathrm{d}\vartheta\, Y_n^m(\vartheta,\varphi)Y_{l'}^{m'}(\vartheta,\varphi)\\ &= \int_0^{2\pi}\mathrm{d}\varphi \int_{-1}^{1}\mathrm{d}\mu\, Y_n^m(\mu,\varphi)Y_{l'}^{m'}(\mu,\varphi) = \delta_{nl'}\delta_{mm'}\end{aligned} \tag{A.4}$$

Similarly, Legendre polynomials form a complete orthogonal system on the interval $[-1, 1]$:

$$\int_{-1}^{1} P_n(\mu)P_m(\mu)\mathrm{d}\mu = \frac{2\delta_{nm}}{2n+1}.$$

Spherical harmonics satisfy the following *addition theorem* (e.g., [49, Remark 3.88]), which allows to express the value of Legendre polynomial of degree k at $\mu_0 = \cos\vartheta_0 = \mathbf{\Omega}\cdot\mathbf{\Omega}'$ (Fig. A.1) as a dot product of vectors with values of the $2n+1$ spherical harmonics at $\mathbf{\Omega}$ and $\mathbf{\Omega}'$, respectively:

$$\begin{aligned}P_n(\mu_0) &= P_n(\mu)P_n(\mu') + 2\sum_{m=1}^{n}\frac{(n-m)!}{(n+m)!}\cos\big(m(\varphi-\varphi')\big)P_n^m(\mu)P_n^m(\mu')\\ &= \frac{4\pi}{2n+1}\sum_{m=-n}^{n} Y_n^m(\vartheta,\varphi)Y_n^m(\vartheta',\varphi'),\end{aligned} \tag{A.5}$$

Note that this greatly simplifies integrals of type

$$\int_{\mathcal{S}_2} P_n(\mathbf{\Omega}\cdot\mathbf{\Omega}')f(\mathbf{\Omega}')\,\mathrm{d}\mathbf{\Omega}'$$

(as in the proof of Lemma 3), because of the complicated form of $\mathbf{\Omega}\cdot\mathbf{\Omega}'$:

$$\begin{aligned}\mathbf{\Omega}\cdot\mathbf{\Omega}' &= [\sin\vartheta\cos\varphi, \sin\vartheta\sin\varphi, \cos\vartheta]^T\cdot[\sin\vartheta'\cos\varphi', \sin\vartheta'\sin\varphi', \cos\vartheta']\\ &= \mu'\mu + \sqrt{(1-\mu'^2)(1-\mu^2)}\cos(\varphi'-\varphi),\end{aligned}$$

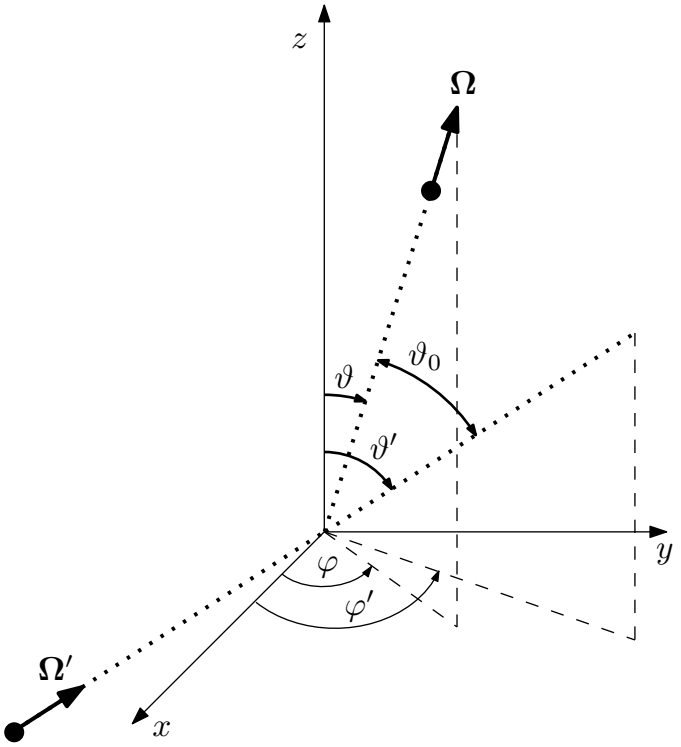


Figure A.1: Geometry of scattering.

Linear combination of spherical harmonics of degree n produces a *surface spherical harmonic of degree n*:

$$\mathcal{Y}_n(\mathbf{\Omega}) = \mathcal{Y}_n(\vartheta, \varphi) =$$
$$A_0 P_n(\cos\vartheta) + \sum_{m=1}^{n} [A_m \cos(m\varphi) P_n^m(\cos\vartheta) + B_m \sin(m\varphi) P_n^m(\cos\vartheta)]$$

Surface spherical harmonics are formally defined as restrictions of homogeneous harmonic polynomials of degree n to unit sphere $\mathcal{S}_2$ ([17, Art. 110], [49, Def. 3.22]).

Appendix B

P_3 advection matrices

In this appendix, we investigate the advection matrices $\mathbf{A}^s_{P_N}$ $(s = x, y, z)$ for the special case $N = 3$ (the statements that follow have been computationally verified to hold for $N = 1, 2, \ldots, 11$ using the symbolic system Mathematica 9.0). It is more convenient for this analysis to consider the steady-state P_N equations (B.1) as a limit of the time-dependent equations

$$\widehat{\mathcal{I}}_{P_N}\left(\frac{\partial}{\partial t} + A + \Sigma_t - K\right)\mathcal{I}_{P_N}\Phi = \widehat{\mathcal{I}}_{P_N} q, \tag{B.1}$$

in which $\frac{\partial \psi}{\partial t} \to 0$, for time-independent boundary conditions and sources $q(\cdot, \cdot, t) = \text{const}$. Then, the advection matrices describe advection of neutrons introduced into the system by the boundary and internal sources, which is in the steady-state limit perfectly balanced by their attenuation due to net effect of collisions of all types.

B. P_3 ADVECTION MATRICES

$$\mathbf{A}^x_{P_N} = \begin{bmatrix}
0 & 0 & 0 & \frac{1}{\sqrt{3}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{1}{\sqrt{5}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{5}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{1}{\sqrt{3}} & 0 & 0 & 0 & 0 & 0 & -\frac{1}{\sqrt{15}} & 0 & \frac{1}{\sqrt{5}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{1}{\sqrt{5}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sqrt{\frac{3}{14}} & 0 & -\frac{1}{\sqrt{70}} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{7}} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -\frac{1}{\sqrt{15}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sqrt{\frac{6}{35}} & 0 & 0 \\
0 & 0 & \frac{1}{\sqrt{5}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\sqrt{\frac{3}{35}} & 0 & \frac{1}{\sqrt{7}} & 0 \\
0 & 0 & 0 & \frac{1}{\sqrt{5}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{\sqrt{70}} & 0 & \sqrt{\frac{3}{14}} \\
0 & 0 & 0 & 0 & \sqrt{\frac{3}{14}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{7}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -\frac{1}{\sqrt{70}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -\sqrt{\frac{3}{35}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \sqrt{\frac{6}{35}} & 0 & -\frac{1}{\sqrt{70}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{7}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sqrt{\frac{3}{14}} & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}$$

$$\mathbf{A}^y_{P_N} = \begin{bmatrix}
0 & \frac{1}{\sqrt{3}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{1}{\sqrt{3}} & 0 & 0 & 0 & 0 & 0 & -\frac{1}{\sqrt{15}} & 0 & -\frac{1}{\sqrt{5}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{5}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{1}{\sqrt{5}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \frac{1}{\sqrt{5}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{\sqrt{70}} & 0 & -\sqrt{\frac{3}{14}} \\
0 & 0 & \frac{1}{\sqrt{5}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\sqrt{\frac{3}{35}} & 0 & -\frac{1}{\sqrt{7}} & 0 \\
0 & -\frac{1}{\sqrt{15}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sqrt{\frac{6}{35}} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{7}} & 0 & 0 & 0 & 0 & 0 \\
0 & -\frac{1}{\sqrt{5}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sqrt{\frac{3}{14}} & 0 & \frac{1}{\sqrt{70}} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sqrt{\frac{3}{14}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{7}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \sqrt{\frac{6}{35}} & 0 & \frac{1}{\sqrt{70}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -\sqrt{\frac{3}{35}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -\frac{1}{\sqrt{70}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -\frac{1}{\sqrt{7}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -\sqrt{\frac{3}{14}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}$$

$$\mathbf{A}^z_{P_N} = \left[\begin{array}{cccccccccccccccc}
0 & 0 & \frac{1}{\sqrt{3}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{5}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{1}{\sqrt{3}} & 0 & 0 & 0 & 0 & 0 & \frac{2}{\sqrt{15}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{5}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{7}} & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{1}{\sqrt{5}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2\sqrt{\frac{2}{35}} & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{2}{\sqrt{15}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{3}{\sqrt{35}} & 0 & 0 & 0 \\
0 & 0 & 0 & \frac{1}{\sqrt{5}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2\sqrt{\frac{2}{35}} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{7}} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{1}{\sqrt{7}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 2\sqrt{\frac{2}{35}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \frac{3}{\sqrt{35}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 2\sqrt{\frac{2}{35}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{7}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}\right]$$

By computing the norms of matrices

$$D_{xy} := \left(\mathbf{A}^x_{P_N}\right)^T \mathbf{A}^y_{P_N} - \left(\mathbf{A}^y_{P_N}\right)^T \mathbf{A}^x_{P_N}$$

(similarly for the remaining combinations of x,y), i.e. (for the largest singular value norm)

$$\|D_{xy}\| = \|D_{yz}\| = \|D_{xz}\| = \frac{2}{5}$$

we observe that the matrices $A^s_{P_3}$ ($s = x, y, z$) do not commute and hence can not be simultaneously diagonalized by a common eigenvector matrix. Consequently, the radiation advected by these matrices cannot be decomposed into plane-waves propagating in distinct directions (as in the case of the S_N approximation), but rather consists of a combination of waves propagating in the infinitely many directions in $\mathbb{R}^3$.

Let us now take an arbitrary fixed direction $\mathbf{n} = [n_x, n_y, n_z]$ from this infinite set. The matrix

$$\mathbf{A}^{\mathbf{n}}_{P_N} = n_x \mathbf{A}^x_{P_N} + n_y \mathbf{A}^y_{P_N} + n_z \mathbf{A}^z_{P_N},$$

displayed below for the case $N = 3$, shows that at most 7 unknowns are coupled in the P_3 system, as the capture matrix $C = \Sigma_t - K$ is diagonal (Corollary 2 on p. 45).

$$\begin{pmatrix}
0 & \frac{n_y}{\sqrt{3}} & \frac{n_z}{\sqrt{3}} & \frac{n_x}{\sqrt{3}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{n_y}{\sqrt{3}} & 0 & 0 & 0 & \frac{n_x}{\sqrt{5}} & \frac{n_z}{\sqrt{5}} & -\frac{n_y}{\sqrt{15}} & 0 & -\frac{n_y}{\sqrt{5}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{n_z}{\sqrt{3}} & 0 & 0 & 0 & 0 & \frac{n_y}{\sqrt{5}} & \frac{2\,n_z}{\sqrt{15}} & \frac{n_x}{\sqrt{5}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{n_x}{\sqrt{3}} & 0 & 0 & 0 & \frac{n_y}{\sqrt{5}} & 0 & -\frac{n_x}{\sqrt{15}} & \frac{n_z}{\sqrt{5}} & \frac{n_x}{\sqrt{5}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{n_x}{\sqrt{5}} & 0 & \frac{n_y}{\sqrt{5}} & 0 & 0 & 0 & 0 & 0 & \sqrt{\frac{3}{14}}\,n_x & \frac{n_z}{\sqrt{7}} & -\frac{n_x}{\sqrt{70}} & 0 & -\frac{n_y}{\sqrt{70}} & 0 & -\sqrt{\frac{3}{14}}\,n_y \\
0 & \frac{n_z}{\sqrt{5}} & \frac{n_y}{\sqrt{5}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{n_x}{\sqrt{7}} & 2\sqrt{\frac{2}{35}}\,n_z & -\sqrt{\frac{3}{35}}\,n_y & 0 & -\frac{n_y}{\sqrt{7}} & 0 \\
0 & -\frac{n_y}{\sqrt{15}} & \frac{2\,n_z}{\sqrt{15}} & -\frac{n_x}{\sqrt{15}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sqrt{\frac{6}{35}}\,n_y & \frac{3\,n_z}{\sqrt{35}} & \sqrt{\frac{6}{35}}\,n_x & 0 & 0 \\
0 & 0 & \frac{n_x}{\sqrt{5}} & \frac{n_z}{\sqrt{5}} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{n_y}{\sqrt{7}} & 0 & -\sqrt{\frac{3}{35}}\,n_x & 2\sqrt{\frac{2}{35}}\,n_z & \frac{n_x}{\sqrt{7}} & 0 \\
0 & -\frac{n_y}{\sqrt{5}} & 0 & \frac{n_x}{\sqrt{5}} & 0 & 0 & 0 & 0 & 0 & \sqrt{\frac{3}{14}}\,n_y & 0 & \frac{n_y}{\sqrt{70}} & 0 & -\frac{n_x}{\sqrt{70}} & \frac{n_z}{\sqrt{7}} & \sqrt{\frac{3}{14}}\,n_x \\
0 & 0 & 0 & 0 & \sqrt{\frac{3}{14}}\,n_x & 0 & 0 & 0 & \sqrt{\frac{3}{14}}\,n_y & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{n_z}{\sqrt{7}} & \frac{n_x}{\sqrt{7}} & 0 & \frac{n_y}{\sqrt{7}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -\frac{n_x}{\sqrt{70}} & 2\sqrt{\frac{2}{35}}\,n_z & \sqrt{\frac{6}{35}}\,n_y & 0 & \frac{n_y}{\sqrt{70}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -\sqrt{\frac{3}{35}}\,n_y & \frac{3\,n_z}{\sqrt{35}} & -\sqrt{\frac{3}{35}}\,n_x & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -\frac{n_y}{\sqrt{70}} & 0 & \sqrt{\frac{6}{35}}\,n_x & 2\sqrt{\frac{2}{35}}\,n_z & -\frac{n_x}{\sqrt{70}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -\frac{n_y}{\sqrt{7}} & 0 & \frac{n_x}{\sqrt{7}} & \frac{n_z}{\sqrt{7}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -\sqrt{\frac{3}{14}}\,n_y & 0 & 0 & 0 & \sqrt{\frac{3}{14}}\,n_x & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}$$

Figure B.1: $\mathbf{A}^{\mathbf{n}}_{\mathrm{P}_3}$

Its eigendecomposition shows that speed of propagation is uniform for all $\mathbf{n} \in \mathbb{R}^3$, given by the eigenvalues corresponding to the case $\|\mathbf{n}\| = 1$ (written with their multiplicities):

$$\left\{0, 0, 0, 0, -\sqrt{\tfrac{3}{7}}, -\sqrt{\tfrac{3}{7}}, \sqrt{\tfrac{3}{7}}, \sqrt{\tfrac{3}{7}}, -\tfrac{1}{\sqrt{7}}, -\tfrac{1}{\sqrt{7}}, \tfrac{1}{\sqrt{7}}, \tfrac{1}{\sqrt{7}},\right.$$
$$\left. -\sqrt{\tfrac{1}{35}\left(15 - 2\sqrt{30}\right)}, \sqrt{\tfrac{1}{35}\left(15 - 2\sqrt{30}\right)}, -\sqrt{\tfrac{1}{35}\left(15 + 2\sqrt{30}\right)}, \sqrt{\tfrac{1}{35}\left(15 + 2\sqrt{30}\right)}\right\}$$

REMARK B.1 (TIME DEPENDENT PROBLEMS). Let us compare the nullspaces of P_2 advection matrix $\mathbf{A}^{\mathbf{n}}_{P_2}$:

$$\begin{bmatrix} \frac{\sqrt{\frac{3}{5}}\left(n_z^2-1\right)}{1-2n_y^2} & \frac{2\sqrt{\frac{3}{5}}n_x n_z}{1-2n_y^2} & \frac{\frac{3n_z^2}{2n_y^2-1}+1}{\sqrt{5}} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \frac{n_x\left(n_z^2-2n_y^2\right)}{n_y\left(2n_y^2-1\right)} & -\frac{n_z-2n_z^3}{n_y-2n_y^3} & \frac{\sqrt{3}n_x n_z^2}{n_y-2n_y^3} \\ -\frac{n_z-n_z^3}{n_y-2n_y^3} & \frac{n_x\left(-2n_y^2-2n_z^2+1\right)}{n_y\left(2n_y^2-1\right)} & \frac{\sqrt{3}n_z\left(2n_y^2+n_z^2-1\right)}{n_y\left(2n_y^2-1\right)} \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

and P_3 advection matrix $\mathbf{A}^{\mathbf{n}}_{P_3}$:

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ \frac{\sqrt{\frac{15}{14}}n_x(n_z^2-1)}{n_y\left(4n_y^2-3\right)} & \frac{\sqrt{\frac{5}{7}}n_z\left(2n_y^2+3n_z^2-3\right)}{n_y\left(4n_y^2-3\right)} & \frac{n_x\left(-4n_y^2-15n_z^2+3\right)}{\sqrt{14}n_y\left(4n_y^2-3\right)} & \frac{\sqrt{\frac{3}{7}}n_z\left(-4n_y^2-5n_z^2+3\right)}{n_y\left(4n_y^2-3\right)} \\ -\frac{\sqrt{\frac{30}{7}}n_x n_z(n_z^2-1)}{8n_y^4-10n_y^2+3} & -\frac{\sqrt{\frac{5}{7}}(2n_z^2-1)\left(4n_y^2+3n_z^2-3\right)}{8n_y^4-10n_y^2+3} & \frac{5\sqrt{\frac{2}{7}}n_x\left(n_y^2-3n_x^2\right)n_z}{8n_y^4-10n_y^2+3} & \frac{\sqrt{\frac{3}{7}}\left(8n_y^4-10n_y^2+10n_z^4+5\left(4n_y^2-3\right)n_z^2+3\right)}{8n_y^4-10n_y^2+3} \\ -\frac{\sqrt{\frac{15}{14}}\left(2n_y^2-2n_z^2-1\right)(n_z^2-1)}{8n_y^4-10n_y^2+3} & \frac{2\sqrt{\frac{5}{7}}n_x n_z\left(2n_y^2-3n_z^2\right)}{8n_y^4-10n_y^2+3} & \frac{8n_y^4-10(n_z^2+1)n_y^2-30n_z^4+15n_z^2+3}{\sqrt{14}\left(8n_y^4-10n_y^2+3\right)} & \frac{10\sqrt{\frac{3}{7}}n_x n_z^3}{8n_y^4-10n_y^2+3} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{n_x\left(-8n_y^4+(4n_z^2+6)n_y^2-3n_z^4+n_z^2-1\right)}{n_y\left(8n_y^4-10n_y^2+3\right)} & \frac{\sqrt{\frac{3}{2}}n_z(-6n_z^4+5n_z^2-1)}{n_y\left(8n_y^4-10n_y^2+3\right)} & \frac{\sqrt{15}n_x n_z^2(3n_z^2-1)}{n_y\left(8n_y^4-10n_y^2+3\right)} & \frac{\sqrt{\frac{5}{2}}n_z^3\left(4n_y^2+6n_z^2-5\right)}{n_y\left(8n_y^4-10n_y^2+3\right)} \\ \frac{\sqrt{\frac{3}{2}}n_z(-4n_z^4+5n_z^2-1)}{n_y\left(8n_y^4-10n_y^2+3\right)} & -\frac{n_x\left(2n_y^2-3n_z^2\right)\left(4n_y^2+4n_z^2-3\right)}{n_y\left(8n_y^4-10n_y^2+3\right)} & \frac{\sqrt{\frac{5}{2}}n_z\left(4n_y^2+3n_z^2-3\right)(4n_z^2-1)}{n_y\left(8n_y^4-10n_y^2+3\right)} & \frac{\sqrt{15}n_x n_z^2\left(-4n_y^2-4n_z^2+3\right)}{n_y\left(8n_y^4-10n_y^2+3\right)} \\ \frac{\sqrt{15}n_x n_z^2(n_z^2-1)}{n_y\left(8n_y^4-10n_y^2+3\right)} & \frac{\sqrt{\frac{5}{2}}n_z(2n_z^2-1)\left(4n_y^2+3n_z^2-3\right)}{n_y\left(8n_y^4-10n_y^2+3\right)} & -\frac{n_x\left(8n_y^4-10n_y^2+15n_z^4+5\left(4n_y^2-3\right)n_z^2+3\right)}{n_y\left(8n_y^4-10n_y^2+3\right)} & -\frac{\sqrt{\frac{3}{2}}n_z\left(16n_y^4+20(n_z^2-1)n_y^2+10n_z^4-15n_z^2+6\right)}{n_y\left(8n_y^4-10n_y^2+3\right)} \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

Unlike the P_3 approximation, we can see that the P_2 approximation contains in its advection nullspace nonzero components of the 0-th moment of angular flux, which is proportional to the scalar flux (total spatial neutron density). Therefore, as a consequence of P_2 approximation, not all scalar flux components are

propagated by the action $\mathbf{A}^{\mathbf{n}}_{P_2}\Psi$. It turns out that this is true for any even-order P_N approximation, which is "probably the most salient argument why even-order expansions should be shunned for time dependent problems" [82, p. 20].

Appendix C

$\mathbf{SP}_N$ matrices

Generated by the Mathematica script available from `https://raw.githubusercontent.com/mhanus/hermes/SN-adaptive/hermes2d/examples/neutronics/SPn-BC.nb`

C.1 N = 5

$$\Phi^s = [\phi_0^s, \phi_2^s, \phi_4^s]^T, \quad \mathrm{Q}^s = q_0[1, -\tfrac{2}{3}, \tfrac{8}{15}]^T$$

$$\mathbf{D}^s = \mathrm{diag}\left\{\frac{1}{3\Sigma_1}, \frac{1}{7\Sigma_3}, \frac{1}{11\Sigma_5}\right\}, \quad \mathbf{G}^s = \begin{bmatrix} 1 & -\frac{1}{4} & \frac{1}{8} \\ -\frac{1}{4} & \frac{7}{12} & -\frac{41}{192} \\ \frac{1}{8} & -\frac{41}{192} & \frac{407}{960} \end{bmatrix}$$

$$\mathbf{C}_t^s = \begin{bmatrix} 1 & -\frac{2}{3} & \frac{8}{15} \\ -\frac{2}{3} & 1 & -\frac{4}{5} \\ \frac{8}{15} & -\frac{4}{5} & 1 \end{bmatrix}, \quad \{\lambda_n\} \doteq \{2.33928, 0.484122, 0.176601\}$$

$$\mathbf{C}_0^s = \begin{bmatrix} 1 & -\frac{2}{3} & \frac{8}{15} \\ -\frac{2}{3} & \frac{4}{9} & -\frac{16}{45} \\ \frac{8}{15} & -\frac{16}{45} & \frac{64}{225} \end{bmatrix}, \quad \mathbf{C}_2^s = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{5}{9} & -\frac{4}{9} \\ 0 & -\frac{4}{9} & \frac{16}{45} \end{bmatrix}, \quad \mathbf{C}_4^s = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{9}{25} \end{bmatrix}$$

C.2 N = 7

$$\Phi^s = [\phi^s_0, \phi^s_2, \phi^s_4, \phi^s_6]^T, \quad \mathrm{Q}^s = q_0[1, -\tfrac{2}{3}, \tfrac{8}{15}, -\tfrac{16}{35}]^T$$

$$\mathbf{D}^s = \mathrm{diag}\left\{\frac{1}{3\Sigma_1}, \frac{1}{7\Sigma_3}, \frac{1}{11\Sigma_5}, \frac{1}{15\Sigma_7}\right\}, \quad \mathbf{G}^s = \begin{bmatrix} 1 & -\frac{1}{4} & \frac{1}{8} & -\frac{5}{64} \\ -\frac{1}{4} & \frac{7}{12} & -\frac{41}{192} & \frac{1}{8} \\ \frac{1}{8} & -\frac{41}{192} & \frac{407}{960} & -\frac{233}{1280} \\ -\frac{5}{64} & \frac{1}{8} & -\frac{233}{1280} & \frac{3023}{8960} \end{bmatrix}$$

$$\mathbf{C}^s_t = \begin{bmatrix} 1 & -\frac{2}{3} & \frac{8}{15} & -\frac{16}{35} \\ -\frac{2}{3} & 1 & -\frac{4}{5} & \frac{24}{35} \\ \frac{8}{15} & -\frac{4}{5} & 1 & -\frac{6}{7} \\ -\frac{16}{35} & \frac{24}{35} & -\frac{6}{7} & 1 \end{bmatrix}, \quad \{\lambda_n\} \doteq \{3.01751, 0.621796, 0.245502, 0.115193\}$$

$$\mathbf{C}^s_0 = \begin{bmatrix} 1 & -\frac{2}{3} & \frac{8}{15} & -\frac{16}{35} \\ -\frac{2}{3} & \frac{4}{9} & -\frac{16}{45} & \frac{32}{105} \\ \frac{8}{15} & -\frac{16}{45} & \frac{64}{225} & -\frac{128}{525} \\ -\frac{16}{35} & \frac{32}{105} & -\frac{128}{525} & \frac{256}{1225} \end{bmatrix}, \quad \mathbf{C}^s_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{5}{9} & -\frac{4}{9} & \frac{8}{21} \\ 0 & -\frac{4}{9} & \frac{16}{45} & -\frac{32}{105} \\ 0 & \frac{8}{21} & -\frac{32}{105} & \frac{64}{245} \end{bmatrix},$$

$$\mathbf{C}^s_4 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{9}{25} & -\frac{54}{175} \\ 0 & 0 & -\frac{54}{175} & \frac{324}{1225} \end{bmatrix}, \quad \mathbf{C}^s_6 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{13}{49} \end{bmatrix}.$$

Appendix D

MCP_3 advection matrices

Unlike in the P_3 case, the advection matrices of the MCP_3 formulation are non-symmetric. However, the matrix

$$\mathbf{A}^{\mathbf{n}}_{\mathrm{MCP}_3} = n_x \mathbf{A}^{x}_{\mathrm{MCP}_3} + n_y \mathbf{A}^{y}_{\mathrm{MCP}_3} + n_z \mathbf{A}^{z}_{\mathrm{MCP}_3}$$

is diagonalizable with real eigenvalues that depend only on the length of $\mathbf{n} \in \mathbb{R}^3$; moreover, the non-zero eigenvalues are exactly the same as in the P_3 case, provided that traceless tensorial moments $\psi^{(n)}$ are being advected (e.g. by use of the detracer operator (5.7) on general symmetric tensors $\hat{\psi^{(n)}}$).

The matrices shown below correspond to the following ordering of unknowns.

$$\phi,$$

$$\psi^{(1)}_1, \psi^{(1)}_2, \psi^{(1)}_3,$$

$$\psi^{(2)}_{11}, \psi^{(2)}_{12}, \psi^{(2)}_{13}, \psi^{(2)}_{22}, \psi^{(2)}_{23}, \psi^{(2)}_{33},$$

$$\psi^{(3)}_{111}, \psi^{(3)}_{211}, \psi^{(3)}_{311}, \psi^{(3)}_{221}, \psi^{(3)}_{321}, \psi^{(3)}_{331}, \psi^{(3)}_{222}, \psi^{(3)}_{322}, \psi^{(3)}_{332}, \psi^{(3)}_{333}.$$

D. MCP_3 ADVECTION MATRICES

$$\mathbf{A}^{x}_{\mathrm{MCP}_3} = \begin{bmatrix}
0 & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & \frac{4}{15} & 0 & 0 & -\frac{2}{15} & 0 & -\frac{2}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{2}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \frac{2}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{2}{3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{6}{35} & 0 & 0 & -\frac{9}{35} & 0 & -\frac{9}{35} & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{12}{35} & 0 & 0 & 0 & 0 & -\frac{3}{35} & 0 & -\frac{3}{35} & 0 \\
0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{12}{35} & 0 & 0 & 0 & 0 & -\frac{3}{35} & 0 & -\frac{3}{35} \\
0 & -\frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{3}{35} & 0 & 0 & \frac{12}{35} & 0 & -\frac{3}{35} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{3}{7} & 0 & 0 & 0 & 0 & 0 \\
0 & -\frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{3}{35} & 0 & 0 & -\frac{3}{35} & 0 & \frac{12}{35} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{2}{5} & 0 & 0 & -\frac{1}{5} & 0 & -\frac{1}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{8}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \frac{8}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -\frac{1}{5} & 0 & 0 & \frac{4}{15} & 0 & -\frac{1}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -\frac{1}{5} & 0 & 0 & -\frac{1}{15} & 0 & \frac{4}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -\frac{2}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -\frac{2}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -\frac{2}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -\frac{2}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}$$

$$\mathbf{A}^{y}_{\mathrm{MCP}_3} = \begin{bmatrix}
0 & 0 & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{2}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & -\frac{2}{15} & 0 & 0 & \frac{4}{15} & 0 & -\frac{2}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{2}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -\frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{12}{35} & 0 & 0 & 0 & 0 & -\frac{3}{35} & 0 & -\frac{3}{35} & 0 \\
0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{3}{35} & 0 & 0 & \frac{12}{35} & 0 & -\frac{3}{35} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{3}{7} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{2}{3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{9}{35} & 0 & 0 & 0 & 0 & \frac{6}{35} & 0 & -\frac{9}{35} & 0 \\
0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{3}{35} & 0 & 0 & 0 & 0 & \frac{12}{35} & 0 & -\frac{3}{35} \\
0 & 0 & -\frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{3}{35} & 0 & 0 & 0 & 0 & -\frac{3}{35} & 0 & \frac{12}{35} & 0 \\
0 & 0 & 0 & 0 & 0 & -\frac{2}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{4}{15} & 0 & 0 & -\frac{1}{5} & 0 & -\frac{1}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{2}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{8}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -\frac{2}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -\frac{1}{5} & 0 & 0 & \frac{2}{5} & 0 & -\frac{1}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{8}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -\frac{1}{15} & 0 & 0 & -\frac{1}{5} & 0 & \frac{4}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{2}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}$$

$$\mathbf{A}^{z}_{\mathrm{MCP}_3} = \begin{bmatrix}
0 & 0 & 0 & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \frac{2}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{2}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & -\frac{2}{15} & 0 & 0 & -\frac{2}{15} & 0 & \frac{4}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -\frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{12}{35} & 0 & 0 & 0 & 0 & -\frac{3}{35} & 0 & -\frac{3}{35} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{3}{7} & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{3}{35} & 0 & 0 & -\frac{3}{35} & 0 & \frac{12}{35} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -\frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{3}{35} & 0 & 0 & 0 & 0 & \frac{12}{35} & 0 & -\frac{3}{35} \\
0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{3}{35} & 0 & 0 & 0 & 0 & -\frac{3}{35} & 0 & \frac{12}{35} & 0 \\
0 & 0 & 0 & \frac{2}{3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{9}{35} & 0 & 0 & 0 & 0 & -\frac{9}{35} & 0 & \frac{6}{35} \\
0 & 0 & 0 & 0 & 0 & 0 & -\frac{2}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{2}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{4}{15} & 0 & 0 & -\frac{1}{15} & 0 & -\frac{1}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -\frac{2}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \frac{8}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{2}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -\frac{1}{15} & 0 & 0 & \frac{4}{15} & 0 & -\frac{1}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{8}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -\frac{1}{5} & 0 & 0 & -\frac{1}{5} & 0 & \frac{2}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}$$

The eigenvalues of $\mathbf{A}^{\mathbf{n}}_{\mathrm{MCP}_3}$ (displayed below) corresponding to $\|\mathbf{n}\| = 1$ are (written with their multiplicities)

$$\left\{0,0,0,0,0,0,0,0,-\sqrt{\tfrac{3}{7}},-\sqrt{\tfrac{3}{7}},\sqrt{\tfrac{3}{7}},\sqrt{\tfrac{3}{7}},-\tfrac{1}{\sqrt{7}},-\tfrac{1}{\sqrt{7}},\tfrac{1}{\sqrt{7}},\tfrac{1}{\sqrt{7}},\right.$$
$$\left.-\sqrt{\tfrac{1}{35}\left(15-2\sqrt{30}\right)},\sqrt{\tfrac{1}{35}\left(15-2\sqrt{30}\right)},-\sqrt{\tfrac{1}{35}\left(15+2\sqrt{30}\right)},\sqrt{\tfrac{1}{35}\left(15+2\sqrt{30}\right)}\right\}$$

D. MCP_3 ADVECTION MATRICES

$$
\begin{pmatrix}
0 & \frac{n_x}{3} & \frac{n_y}{3} & \frac{n_z}{3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
n_x & 0 & 0 & 0 & \frac{4\,n_x}{15} & \frac{2\,n_y}{5} & \frac{2\,n_z}{5} & -\frac{2\,n_x}{15} & 0 & -\frac{2\,n_x}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
n_y & 0 & 0 & 0 & -\frac{2\,n_y}{15} & \frac{2\,n_x}{5} & 0 & \frac{4\,n_y}{15} & \frac{2\,n_z}{5} & -\frac{2\,n_y}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
n_z & 0 & 0 & 0 & -\frac{2\,n_z}{15} & 0 & \frac{2\,n_x}{5} & -\frac{2\,n_z}{15} & \frac{2\,n_y}{5} & \frac{4\,n_z}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{2\,n_x}{3} & -\frac{n_y}{3} & -\frac{n_z}{3} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{6\,n_x}{35} & \frac{12\,n_y}{35} & \frac{12\,n_z}{35} & -\frac{9\,n_x}{35} & 0 & -\frac{9\,n_x}{35} & -\frac{3\,n_y}{35} & -\frac{3\,n_z}{35} & -\frac{3\,n_y}{35} & -\frac{3\,n_z}{35} \\
0 & \frac{n_y}{2} & \frac{n_x}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{3\,n_y}{35} & \frac{12\,n_x}{35} & 0 & \frac{12\,n_y}{35} & \frac{3\,n_z}{7} & -\frac{3\,n_y}{35} & -\frac{3\,n_x}{35} & 0 & -\frac{3\,n_x}{35} & 0 \\
0 & \frac{n_z}{2} & 0 & \frac{n_x}{2} & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{3\,n_z}{35} & 0 & \frac{12\,n_x}{35} & -\frac{3\,n_z}{35} & \frac{3\,n_y}{7} & \frac{12\,n_z}{35} & 0 & -\frac{3\,n_x}{35} & 0 & -\frac{3\,n_x}{35} \\
0 & -\frac{n_x}{3} & \frac{2\,n_y}{3} & -\frac{n_z}{3} & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{3\,n_x}{35} & -\frac{9\,n_y}{35} & -\frac{3\,n_z}{35} & \frac{12\,n_x}{35} & 0 & -\frac{3\,n_x}{35} & \frac{6\,n_y}{35} & \frac{12\,n_z}{35} & -\frac{9\,n_y}{35} & -\frac{3\,n_z}{35} \\
0 & 0 & \frac{n_z}{2} & \frac{n_y}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{3\,n_z}{35} & -\frac{3\,n_y}{35} & 0 & \frac{3\,n_x}{7} & 0 & -\frac{3\,n_z}{35} & \frac{12\,n_y}{35} & \frac{12\,n_z}{35} & -\frac{3\,n_y}{35} \\
0 & -\frac{n_x}{3} & -\frac{n_y}{3} & \frac{2\,n_z}{3} & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{3\,n_x}{35} & -\frac{3\,n_y}{35} & -\frac{9\,n_z}{35} & -\frac{3\,n_x}{35} & 0 & \frac{12\,n_x}{35} & -\frac{3\,n_y}{35} & -\frac{9\,n_z}{35} & \frac{12\,n_y}{35} & \frac{6\,n_z}{35} \\
0 & 0 & 0 & 0 & \frac{2\,n_x}{5} & -\frac{2\,n_y}{5} & -\frac{2\,n_z}{5} & -\frac{n_x}{5} & 0 & -\frac{n_x}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{4\,n_y}{15} & \frac{8\,n_x}{15} & 0 & -\frac{n_y}{5} & -\frac{2\,n_z}{15} & -\frac{n_y}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{4\,n_z}{15} & 0 & \frac{8\,n_x}{15} & -\frac{n_z}{15} & -\frac{2\,n_y}{15} & -\frac{n_z}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -\frac{n_x}{5} & \frac{8\,n_y}{15} & -\frac{2\,n_z}{15} & \frac{4\,n_x}{15} & 0 & -\frac{n_x}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{n_z}{3} & \frac{n_y}{3} & 0 & \frac{n_x}{3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -\frac{n_x}{5} & -\frac{2\,n_y}{15} & \frac{8\,n_z}{15} & -\frac{n_x}{15} & 0 & \frac{4\,n_x}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -\frac{n_y}{5} & -\frac{2\,n_x}{5} & 0 & \frac{2\,n_y}{5} & -\frac{2\,n_z}{5} & -\frac{n_y}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -\frac{n_z}{15} & 0 & -\frac{2\,n_x}{15} & \frac{4\,n_z}{15} & \frac{8\,n_y}{15} & -\frac{n_z}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -\frac{n_y}{15} & -\frac{2\,n_x}{15} & 0 & -\frac{n_y}{5} & \frac{8\,n_z}{15} & \frac{4\,n_y}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -\frac{n_z}{5} & 0 & -\frac{2\,n_x}{5} & -\frac{n_z}{5} & -\frac{2\,n_y}{5} & \frac{2\,n_z}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}
$$

Figure D.1: $\mathbf{A}^{\mathbf{n}}_{\mathrm{MCP}_3}$

Appendix E

Tensor identities

The following tensor identities are used in Chap. 5.

$$\nabla \cdot \mathscr{S}\left(\nabla \otimes \mathbb{A}^{(2)}\right) = \frac{2}{3}\mathscr{S}\left(\nabla \otimes \nabla \cdot \mathbb{A}^{(2)}\right) + \frac{1}{3}\nabla^2 \mathbb{A}^{(2)} \tag{E.1}$$

$$\nabla \cdot \mathscr{S}\left(\nabla \otimes \mathbb{A}^{(1)}\right) = \frac{1}{2}\nabla\nabla \cdot \mathbb{A}^{(1)} + \frac{1}{2}\nabla^2 \mathbb{A}^{(1)} \tag{E.2}$$

$$\nabla \cdot \mathscr{S}\left(\mathbb{I} \otimes \nabla \cdot \mathbb{A}^{(2)}\right) = \frac{2}{3}\mathscr{S}\left(\nabla \otimes \nabla \cdot \mathbb{A}^{(2)}\right) + \frac{1}{3}\mathscr{S}\left(\mathbb{I} \otimes \nabla \cdot \nabla \cdot \mathbb{A}^{(2)}\right) \tag{E.3}$$

where $\mathbb{A}^{(n)}$ is a totally symmetric traceless rank-n tensor.

These identities can be derived using the index notation and definitions of the relevant operators (and verified relatively easily using a computational algebra system like Mathematica). Letting

$$\partial_\alpha = \frac{\partial}{\partial x_\alpha}, \quad \partial^2_{\alpha\beta} = \frac{\partial^2}{\partial x_\alpha \partial x_\beta}$$

to simplify the notation, identity (E.1), for instance, is obtained by using

$$\begin{aligned} T := \nabla \cdot \mathscr{S}\left(\nabla \otimes \mathbb{A}^{(2)}\right) &= \frac{1}{3}\partial_\alpha\left(\partial_\gamma A^{(2)}_{\alpha\beta} + \partial_\beta A^{(2)}_{\alpha\gamma} + \partial_\alpha A^{(2)}_{\beta\gamma}\right) \\ &= \frac{1}{3}\left(\partial^2_{\alpha\gamma} A^{(2)}_{\alpha\beta} + \partial^2_{\alpha\beta} A^{(2)}_{\alpha\gamma} + \partial^2_{\alpha\alpha} A^{(2)}_{\beta\gamma}\right) \end{aligned}$$

in

$$\mathscr{S}\left(\nabla \otimes \nabla \cdot \mathbb{A}^{(2)}\right) = \frac{1}{2}\left(\partial^2_{\alpha\beta} A^{(2)}_{\beta\gamma} + \partial^2_{\gamma\beta} A^{(2)}_{\beta\alpha}\right) = \frac{3}{2}T - \frac{1}{2}\partial^2_{\alpha\alpha} A^{(2)}_{\beta\gamma} = \frac{3}{2}T - \frac{1}{2}\nabla^2 \otimes \mathbb{A}^{(2)}$$

so that (dropping the $\otimes$ sign from the last term as ∇^2 is a scalar operator)

$$T = \frac{2}{3}\mathscr{S}\left(\nabla \otimes \nabla \cdot \mathbb{A}^{(2)}\right) + \frac{1}{3}\nabla^2 \mathbb{A}^{(2)}.$$

Identity (E.3) is a special case of the general identity [27, B.4].

Appendix F

On the origin of smoothed aggregations

A short note dedicated to prof. Karel Segeth at the occasion of his 70th birthday.

Authors:

Pavla Fraňková, Milan Hanuš, Hana Kopincová, Roman Kužel, Petr Vaněk and Zbyněk Vastl

F.1 Introduction

The smoothed aggregation method [112, 113, 114, 116] proved to be a very efficient tool for solving various types of elliptic problems and their singular perturbations. In this short note, we turn to the very roots of smoothed aggregation method and derive its two-level variant on a systematic basis.

The multilevel method consists in combination of a coarse-grid correction and smoothing. The coarse-grid correction of a standard two-level method is derived using the A-orthogonal projection of an error to the range of the prolongator. In other words, the coarse-grid correction vector is chosen to minimize the error *after coarse-grid correction procedure*. This means, the standard two-level method minimizes the error in an intermediate stage of the iteration, while we are, naturally,

interested in minimizing *the final error after accomplishing the entire iteration*. In other words, we strive to minimize the error after after coarse-grid correction and subsequent smoothing. The two-level smoothed aggregation method is obtained by solving this minimization problem. This, in the opinion of the authors, explains its remarkable robustness.

We derive the two-level smoothed aggregation method from the variational objective to minimize the error after coarse-grid correction and subsequent post-smoothing. Then, by a trivial argument, we extend our result to the two-level method with pre-smoothing, coarse-grid correction and post-smoothing.

The minimization of error after coarse-grid correction and subsequent smoothing leads to a method with smoothed prolongator. We can say that by smoothing the prolongator, we adapt the coarse-space (the range of the prolongator) to the post-smoother so that the resulting iteration is as efficient as possible. Our short explanation applies to any two-level method with smoothed prolongator. The particular case we have in mind is, however, a method with smoothed *tentative* prolongator given by generalized unknowns aggregations [116]. The discrete basis functions of the coarse-space (the columns of the prolongator) given by unknowns aggregations have no overlap; the natural overlap of discrete basis functions (like it is in the case of finite element basis functions) is created by smoothing and, for additive point-wise smoothers, leads to sparse coarse-level matrix.

Our argument is basically trivial. It, however, shows a fundamental property of the method with smoothed prolongator, that is essential. This argument is known to the authors for a long time, but has never been published.

We conclude our paper by a numerical test. Namely, we demonstrate experimentally that smoothed aggregation method with powerful smoother and small coarse-space solves efficiently highly anisotropic problems without the need to perform semi-coarsening (the coarsening that follows only strong connections).

F.2 Two-level method

We solve a system of linear algebraic equations

$$A\mathbf{x} = \mathbf{f}, \tag{F.1}$$

where A is a symmetric positive definite matrix of order n and $\mathbf{f} \in \mathbb{R}^n$. We assume that an injective linear *prolongator* $p : \mathbb{R}^m \to \mathbb{R}^n,\ m < n$ is given.

The two-level method consists in the combination of a *coarse-grid correction* and *smoothing*. The smoothing means using point-wise iterative methods at the beginning and at the end of the iteration. The coarse-grid correction is derived by correcting an error $\mathbf{e}$ by a coarse-level vector $\mathbf{v}$ so that the resulting error $\mathbf{e} - p\mathbf{v}$ is minimal in A-norm. In other words, we solve the minimization problem

$$\text{find } \mathbf{v} \in \mathbb{R}^m \text{ so that } \|\mathbf{e} - p\mathbf{v}\|_A \text{ is minimal.} \tag{F.2}$$

It is well-known that such vector $p\mathbf{v}$ is an A-orthogonal projection of the error $\mathbf{e}$ onto Range(p), with the projection operator given by

$$Q = p(p^T A p)^{-1} p^T A.$$

Thus, the error propagation operator of the coarse-grid correction is given by $I - Q = I - p(p^T A p)^{-1} p^T A$ and the error propagation operator of the two-level method by

$$E_{TGM} = S_{post}[I - p(p^T A p)^{-1} p^T A] S_{pre}, \tag{F.3}$$

where S_{pre} and S_{post} are error propagation operators of pre- and post- smoothing iterations, respectively.

Clearly, for the error $\mathbf{e}(\mathbf{x}) \equiv \mathbf{x} - A^{-1}\mathbf{f}$ we have $A\mathbf{e}(\mathbf{x}) = A\mathbf{x} - \mathbf{f}$. Hence, the coarse-grid correction can be algorithmized as

$$\mathbf{x} \leftarrow \mathbf{x} - p(p^T A p)^{-1} p^T (A\mathbf{x} - \mathbf{f})$$

and the variational two-level algorithm with post-smoothing step proceeds as follows:

Algorithm 1.

1. Pre-smooth: $\mathbf{x} \leftarrow \mathcal{S}_{pre}(\mathbf{x}, \mathbf{f})$,
2. evaluate the residual: $\mathbf{d} = A\mathbf{x} - \mathbf{f}$,
3. restrict the residual: $\mathbf{d}_2 = p^T \mathbf{d}$,
4. solve a coarse-level problem $A_2\mathbf{v} = \mathbf{d}_2, \quad A_2 = p^T A p$,

5. correct the approximation $\mathbf{x} = \mathbf{x} - p\mathbf{v}$,

6. post-smooth $\mathbf{x} = \mathcal{S}_{post}(\mathbf{x}, \mathbf{f})$.

Here, $\mathcal{S}_{pre}(.,.)$ and $\mathcal{S}_{post}(.,.)$, respectively, represent one or more iterations of point-wise iterative methods for solving (F.1).

The coarse-grid correction vector $\mathbf{v}$ is chosen to minimize the error after Step 5 of Algorithm 1. Thus, we conclude that in the case of a standard variational multigrid, the coarse-grid correction procedure minimizes the error in an intermediate stage of the iteration, while we are in fact interested in minimizing the final error after accomplishing the entire iteration. This means to minimize the error after coarse-grid correction with subsequent smoothing.

F.3 The smoothed aggregation two-level method

In the smoothed aggregation method, we construct the coarse-grid correction to minimize the error *after coarse-grid correction with subsequent smoothing*, which means the final error on the exit of the iteration procedure. The minimization of the error after pre-smoothing, coarse-grid correction and post-smoothing then follows immediately by a trivial argument.

Let S be the error propagation operator of the post-smoother $\mathcal{S}(.,.) = \mathcal{S}_{post}(.,.)$. Throughout this section we assume that S is sparse. This is due to the fact that the above minimization problem leads to smoothed prolongator $P = Sp$ and we need a sparse coarse-level matrix $A_2 = P^T AP$. The additive point-wise smoothing methods have, in general, sparse error propagation operator; this is the case of Jacobi method or Richardson's iteration.

For a multilevel method with post-smoothing only, the error after coarse-grid correction and subsequent smoothing is given by

$$S(\mathbf{e} - p\mathbf{v}), \tag{F.4}$$

where $\mathbf{v}$ is a correction vector and $\mathbf{e}$ the error on the entry of the iteration procedure. We choose $\mathbf{v}$ so that the error in (F.4) is minimal in A-norm, that is, we solve the minimization problem

$$\text{find } \mathbf{v} \in \mathbb{R}^m \text{ such that } \|S(\mathbf{e} - p\mathbf{v})\|_A \text{ is minimal.} \tag{F.5}$$

Since $\|S(\mathbf{e} - p\mathbf{v})\|_A = \|\mathbf{e} - p\mathbf{v}\|_{S^TAS}$, the minimum is attained for $\mathbf{v}$ satisfying

$$\langle S^TAS(\mathbf{e} - p\mathbf{v}), p\mathbf{w}\rangle = 0 \ \forall \mathbf{w} \in \mathbb{R}^m.$$

We have $\langle S^TAS(\mathbf{e} - p\mathbf{v}), p\mathbf{w}\rangle = \langle p^TS^TAS(\mathbf{e} - p\mathbf{v}), \mathbf{w}\rangle$, hence the above identity is equivalent to $p^TS^TASp\mathbf{v} = p^TS^TAS\mathbf{e}$ and setting $P = Sp$, it becomes

$$P^TAP\mathbf{v} = P^TAS\mathbf{e}. \tag{F.6}$$

Here, $\mathbf{e}$ is the error on the entry of the iteration procedure. Assume for now that P is injective. Then by (F.6), we have $\mathbf{v} = (P^TAP)^{-1}P^TAS\mathbf{e}$ and the error after coarse-grid correction and subsequent smoothing is given by

$$S(\mathbf{e} - p\mathbf{v}) = S\left[\mathbf{e} - p(P^TAP)^{-1}P^TAS\mathbf{e}\right] = \left[I - P(P^TAP)^{-1}P^TA\right]S\mathbf{e}. \tag{F.7}$$

By comparing the operator

$$E = \left[I - P(P^TAP)^{-1}P^TA\right]S \tag{F.8}$$

on the right-hand side of (F.7) with (F.3), we identify E as the error propagation operator of the variational multigrid with smoothed prolongator $P = Sp$ and pre-smoothing step given by $\mathbf{x} \leftarrow \mathcal{S}(\mathbf{x}, \mathbf{f})$. The algorithm is as follows:

Algorithm 2.

1. Pre-smooth: $\mathbf{x} \leftarrow \mathcal{S}(\mathbf{x}, \mathbf{f})$,
2. evaluate the residual: $\mathbf{d} = A\mathbf{x} - \mathbf{f}$,
3. restrict the residual: $\mathbf{d}_2 = P^T\mathbf{d}$,
4. solve the coarse-level problem: $A_2\mathbf{v} = \mathbf{d}_2, \quad A_2 = P^TAP$,
5. correct the approximation: $\mathbf{x} \leftarrow \mathbf{x} - P\mathbf{v}$.

Remark 1. Note that in the process of the deriving the algorithm in (F.7), our post-smoother have become a pre-smoother. Nothing was lost in that process; the algorithm minimizes the final error and takes into account the pre-smoother.

Remark 2. The smoothed prolongator $P = Sp$ is potentially non-injective, hence the coarse-level matrix $A_2 = P^TAP$ is potentially singular. In this case, we need to replace the inverse of P^TAP in (F.7) by a pseudo-inverse.

We summarize our considerations in the form of a theorem.

Theorem 3. *The error propagation operator E in (F.8) (the error propagation operator of Algorithm 2) satisfies the identity*

$$\|E\mathbf{e}\|_A = \inf_{\mathbf{v}\in\mathbb{R}^m} \|S(\mathbf{e} - p\mathbf{v})\|_A$$

for all $\mathbf{e} \in \mathbb{R}^n$.

Proof. The proof follows directly from the fact that Algorithm 2 was derived from variational objective (F.5). □

Remark 4. One may also start with the variational objective to minimize the final error after performing the pre-smoothing, the coarse-grid correction and the post-smoothing. Such extension is trivial, the pre-smoother has no influence on the coarse-grid correction operator $I - P(P^TAP)^{-1}P^TA$ and influences only its argument. Indeed, assuming the error propagation operator of the pre-smoother is S^* (the A-adjoint operator), the final error is given by $S(S^*\mathbf{e} - p\mathbf{v})$ and we solve the minimization problem

$$\text{for } \mathbf{e} \in \mathbb{R}^n \text{ find } \mathbf{v} \in \mathbb{R}^m \ : \ \|S(S^*\mathbf{e} - p\mathbf{v})\|_A \text{ is minimal.} \tag{F.9}$$

Fundamentally, this is the same minimization problem as (F.5); to derive the corresponding algorithm, it is simply sufficient to follow our manipulations from (F.5) to (F.7) with $\mathbf{e} \leftarrow S^*\mathbf{e}$. This way, we end up with a two-level method that has the error propagation operator

$$E = \left[I - P(P^TAP)^{-1}P^TA\right] SS^*, \tag{F.10}$$

(see (F.3)) that is, with the algorithm

Algorithm 3.

1. Pre-smooth: $\mathbf{x} \leftarrow \mathcal{S}_t(\mathbf{x}, \mathbf{f})$, where $\mathcal{S}_t$ is an iterative method with error propagation operator S^*.

2. pre-smooth: $\mathbf{x} \leftarrow \mathcal{S}(\mathbf{x}, \mathbf{f})$, where $\mathcal{S}$ is an iterative method with error propagation operator S.

3. evaluate the residual: $\mathbf{d} = A\mathbf{x} - \mathbf{f}$,

4. restrict the residual: $\mathbf{d}_2 = P^T\mathbf{d}$,

5. solve the coarse-level problem: $A_2\mathbf{v} = \mathbf{d}_2, \quad A_2 = P^T AP$,

6. correct the approximation: $\mathbf{x} \leftarrow \mathbf{x} - P\mathbf{v}$.

We summarize the content of Remark 4 as a theorem.

Theorem 5. *The error propagation operator (F.10) of Algorithm 3 satisfies the identity*

$$\|E\mathbf{e}\|_A = \inf_{\mathbf{v}\in\mathbb{R}^m} \|S(S^*\mathbf{e} - p\mathbf{v})\|_A$$

for all $\mathbf{e} \in \mathbb{R}^n$.

Proof. The proof follows directly from the fact that Algorithm 3 was derived from variational objective (F.9). □

Remark 6. Our manipulations hold equally for a general pre-smoother with error propagation operator $M \neq S^*$, simply by replacing $S^* \leftarrow M$. The error propagation operator M has no influence on the coarse-space and thus it does not have to be sparse.

F.4 Numerical example

To demonstrate the robustness of smoothed aggregation method, we consider the algorithm that is a modification of the method proposed and analyzed in [115]. Its relationship to Algorithm 2 is obvious. This method uses the smoothing iterative method $\mathcal{S}(\cdot,\cdot)$ which is a sequence of Richardson's iterations with carefully chosen iteration parameters. The error propagation operator S of the smoother $\mathcal{S}(\cdot,\cdot)$ is therefore a polynomial in the matrix A.

In this method, we use massive smoother S and a small coarse-space resulting in sparse coarse-level matrix.

Let $\bar{\lambda} \geq \varrho(A)$ and d be the desired degree of the smoothing polynomial S. We set

$$\alpha_i = \left[\frac{\bar{\lambda}}{2}\left(1 - \cos\frac{2i\pi}{2d+1}\right)\right]^{-1}, \quad i = 1,\ldots,d, \tag{F.11}$$

$$S = (I - \alpha_1 A)\ldots(I - \alpha_d A) \tag{F.12}$$

and

$$P = Sp.$$

Here, p is a *tentative prolongator* given by generalized unknowns aggregation. The simplest aggregation method is described in this section.

The smoother S is chosen to minimize $\varrho(S^2A)$. The reason for this comes from the fact that the convergence of the method of [115] is guided by the constant C in the weak approximation condition

$$\forall \mathbf{e} \in \mathbb{R}^n \; \exists \mathbf{v} \in \mathbb{R}^m \; : \; \|\mathbf{e} - p\mathbf{v}\| \leq \frac{C}{\sqrt{\varrho(S^2A)}} \|\mathbf{e}\|_A. \tag{F.13}$$

The smaller $\varrho(S^2A)$, the easier it becomes to satisfy (F.13) with a reasonable (sufficiently small) constant. It holds that ([115])

$$\bar{\lambda}_{S^2A} \equiv \frac{\bar{\lambda}}{(1+2d)^2} \geq \varrho(S^2A). \tag{F.14}$$

The aggregates $\{\mathcal{A}_j\}$ are sets of fine-level degrees of freedom that form a disjoint covering of the set of all fine-level degrees of freedom. For example, we can choose aggregates to form a decomposition of the set of degrees of freedom induced by a geometrically reasonable partitioning of the computational domain. For standard discretizations of scalar elliptic problems, the tentative prolongator matrix p is the $n \times m$ matrix ($m =$ the number of aggregates)

$$p_{ij} = \begin{cases} 1 & \text{if } \; i \in \mathcal{A}_j \\ 0 & \text{otherwise} \end{cases} \tag{F.15}$$

that is, the j-th column is created by restricting a vector of ones onto the j-th aggregate, with zeroes elsewhere. Thus, the aggregation method can be viewed as a piece-wise constant coarsening in a discrete sense. The generalized aggregation method, suitable for non-scalar elliptic problems (like that of linear elasticity), is described in [116].

Algorithm 4. Given the degree d of the smoothing polynomial $S = pol(A)$, the smoothed prolongator $P = Sp$ where p is the tentative prolongator and the prolongator smoother S is given by (F.12), the upper bound $\bar{\lambda} \geq \varrho(A)$ and a parameter $\omega \in (0, 1)$, one iteration of the two-level algorithm

$$\mathbf{x} \leftarrow TG(\mathbf{x}, \mathbf{f})$$

proceeds as follows:

1. perform
$$\mathbf{x} \leftarrow \mathbf{x} - \frac{\omega}{\bar{\lambda}_{S^2A}} S^2(A\mathbf{x} - \mathbf{f}),$$
where $\bar{\lambda}_{S^2A}$ is given by (F.14) and S by (F.12),

2. perform the iteration with symmetric error propagation operator S given by (F.12), that is,
for $i = 1, \ldots, d$ do
$$\mathbf{x} \leftarrow (I - \alpha_i A)\,\mathbf{x} + \alpha_i \mathbf{f},$$

3. evaluate the residual $\mathbf{d} = A\mathbf{x} - \mathbf{f}$,

4. restrict the residual $\mathbf{d}_2 = P^T\mathbf{d}$,

5. solve the coarse-level problem $A_2\mathbf{v} = \mathbf{d}_2$, $A_2 = P^TAP$,

6. correct the approximation $\mathbf{x} \leftarrow \mathbf{x} - P\mathbf{v}$,

7. for $i = 1, \ldots, d$ do
$$\mathbf{x} \leftarrow (I - \alpha_i A)\,\mathbf{x} + \alpha_i \mathbf{f},$$

8. perform
$$\mathbf{x} \leftarrow \mathbf{x} - \frac{\omega}{\bar{\lambda}_{S^2A}} S^2(A\mathbf{x} - \mathbf{f}).$$

Thus, Algorithm 4 is a symmetrized version of Algorithm 2 with added smoothing in steps 1 and 8.

It is generally believed that in order to solve efficiently an anisotropic problem, one has to perform coarsening only by following *strong connections*. This technique is called *semi-coarsening*. In our case, we form aggregates by coarsening by a factor of 10 in all 3 spatial directions, which means, we do not perform semi-coarsening. Despite of this fact, our method gives satisfactory results regardless of the anisotropy coefficient ε. In this experiment, the symmetric Algorithm 4 is used as a conjugate gradient method preconditioner.

Test problem

- Problem:
$$-\left(\frac{\partial^2}{\partial x^2} + \varepsilon\frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) u = f \text{ on } \Omega = (0,1)^3,\ u = 0 \text{ on } \partial\Omega. \qquad \text{(F.16)}$$

512 000 dofs, coarse space 512 dofs, $\deg(S) = 7$, $H/h = 9$.		
ε	rate of conv. q_N	no. iter. N
1000	0.321	19
100	0.241	15
10	0.137	11
1	0.131	11
0.1	0.221	14
0.01	0.317	19
0.001	0.300	18

Table F.1: 3D anisotropic problem

- Mesh: $82 \times 82 \times 82$ regular square mesh, 512 000 unconstrained degrees of freedom.
- Aggregates: cubic groups of $10 \times 10 \times 10$ unconstrained vertices.
- Coarse-space size: 512 degrees of freedom.
- Degree of smoothing polynomial: 7.
- Stopping criterion: relative residual $< 10^{-9}$.

The results are summed up in Table F.1. Note that here, the estimate of the rate of convergence after N iterations is defined as

$$q_N = \left(\|A\mathbf{x}^N - \mathbf{f}\|/\|A\mathbf{x}^0 - \mathbf{f}\|\right)^{\frac{1}{N}}.$$

Here, $\mathbf{x}^i$ denotes the i-th iteration.

Acknowledgements

This work was sponsored by the TAČR (Technologická Agentura České Republiky) grant TA01020352, ITI (Institut Teoretické Informatiky) grant 1R0545, Department of the Navy Grant N62909-11-1-7032.

Literature

[1] Ron T. Ackroyd. On a rigorous resolution of the transport equation into a system of diffusion-like equations. *Progress in Nuclear Energy*, 35(1):1–64, 1999. 86

[2] Marvin L. Adams. "I have an idea!" An appreciation of Edward W. Larsen's contributions to particle transport. *Annals of Nuclear Energy*, 31:1963–1986, 2004. 61

[3] Marvin L. Adams and Edward W. Larsen. Fast iterative methods for discrete-ordinates particle transport calculations. *Progress in Nuclear Energy*, 40:3–159, 2002. 56, 62

[4] V. Agoshkov. *Boundary Value Problems for Transport Equations*. Modeling and Simulation in Science, Engineering and Technology. Birkhäuser Boston, 1998. ISBN 9780817639860. 12, 17, 19, 22, 23

[5] Grégoire Allaire and Guillaume Bal. Homogenization of the criticality spectral equation in neutron transport. *RAIRO - Modélisation mathématique et analyse numérique*, 33(4):721–746, 1999. 28

[6] Jon Applequist. Traceless cartesian tensor forms for spherical harmonic functions: new theorems and applications to electrostatics of dielectric media. *Journal of Physics A: Mathematical and General*, 22:4303–4330, 1989. 90, 91, 93, 94, 101

[7] Jon Applequist. Maxwell-Cartesian spherical harmonics in multipole potentials and atomic orbitals. *Theoretical Chemistry Accounts*, 107:103–115, 2002. 3, 90, 91, 92, 96

[8] Mario Arioli, Jörg Liesen, Agnieszka Miedlar, and Zdeněk Strakoš. Interplay between discretization and algebraic computation in adaptive numerical solutionof elliptic pde problems. *GAMM-Mitteilungen*, 36(1):102–129, 2013. ISSN 1522-2608. doi: 10.1002/gamm.201310006. URL `http://dx.doi.org/10.1002/gamm.201310006`. 72

[9] Yousry Azmy and Enrico Sartori. *Nuclear Computational Science. A Century in Review*. Springer, 2006. 23, 62

[10] Satish Balay, Shrirang Abhyankar, Mark F. Adams, Jed Brown, Peter Brune, Kris Buschelman, Victor Eijkhout, William D. Gropp, Dinesh Kaushik, Matthew G. Knepley, Lois Curfman McInnes, Karl Rupp, Barry F. Smith, and Hong Zhang. PETSc Web page. `http://www.mcs.anl.gov/petsc`, 2014. URL `http://www.mcs.anl.gov/petsc`. 6, 147

[11] Mohamed Boulanouar. New trace theorems for neutronic function spaces. *Transport Theory and Statistical Physics*, 38:228–242, 2009. 19

[12] L. Bourhrara. New variational formulations for the neutron transport equation. *Transport Theory and Statistical Physics*, 33(2):93–124, 2004. doi: 10.1081/TT-120037803. 23

[13] L. Bourhrara. H1 approximations of the neutron transport equation and associated diffusion equations. *Transport Theory and Statistical Physics*, 35 (3-4):89–108, 2006. doi: 10.1080/00411450600901730. 5

[14] L. Bourhrara. W n approximations of neutron transport equation. *Transport Theory and Statistical Physics*, 38(4):195–227, 2009. doi: 10.1080/00411450903192938. 23

[15] Patrick S. Brantley and Edward W. Larsen. The simplified P3 approximation. *Nuclear Science and Engineering*, 134:1–21, 2000. 74, 75, 77

[16] Marian Brezina, Petr Vaněk, and Panayot S. Vassilevski. An improved convergence analysis of smoothed aggregation algebraic multigrid. *Numerical Linear Algebra with Applications*, 19(3):441–469, 2012. ISSN 1099-1506. doi: 10.1002/nla.775. URL `http://dx.doi.org/10.1002/nla.775`. 51

[17] William E. Byerly. *An elementary treatise on Fourier's series and spherical, cylindrical and ellipsoidal harmonics, with applications to problems in mathematical physics*. Ginn and Co., Boston, USA, 2nd edition, 1893. 86, 92, 163

[18] Dan Gabriel Cacuci. *Handbook of Nuclear Engineering, Volume I: Nuclear Engineering Fundamentals*. Springer Science + Business Media LLC, 2010. 34

[19] M. Capilla, C. F. Talavera, D. Ginestar, and G. Verdú. A nodal collocation approximation for the multi-dimensional PL equations – 2D applications. *Annals of Nuclear Energy*, 35:1820–1830, 2008. 52, 85

[20] K.M. Case and P.F. Zweifel. *Linear transport theory*. Addison-Wesley series in nuclear engineering. Addison-Wesley Pub. Co., 1967. 29

[21] Y. A. Chao and Y. A. Shatilla. Conformal Mapping and Hexagonal Nodal Methods – II: Implementation in the ANC-H Code. *Nucl. Sci. Eng.*, 121: 210–225, 1995. 136, 138

[22] Jin-Young Cho, Kang-Seog Kim, Chung-Chan Lee, Sung-Quun Zee, and Han-Gyu Joo. Axial SPN and radial MOC coupled whole core transport calculation. *Journal of Nuclear Science and Technology*, 44:1156–1171, 2007. 75

[23] Jin-Young Cho, Kang-Seog Kim, Hyung-Jin Shim, Jae-Seung Song, Chung-Chan Lee, and Han-Gyu Joo. Whole core transport calculation employing hexagonal modular ray tracing and CMFD formulation. *Journal of Nuclear Science and Technology*, 45(8):740–751, 2008. 37

[24] Nam Zin Cho. Fundamentals and recent developments of reactor physics methods. *Nuclear Engineering and Technology*, 37:25–78, 2005. 34, 51

[25] R. Ciolini, G. G. M. Coppa, B. Montagnini, and P. Ravetto. Simplified P_N and A_N Methods in Neutron Transport. *Progress in Nuclear Energy*, 40(2): 237–264, 2002. 73, 86, 124

[26] Chris Clarkson and Bob Osano. Locally extracting scalar, vector and tensor modes in cosmological perturbation theory. *Classical and Quantum Gravity*, 29(7):079601, 2012. URL `http://arxiv.org/pdf/1102.4265.pdf`. 159

[27] G. G. M. Coppa. Deduction of a symmetric tensor formulation of the p_n method for the linear transport equation. *Progress in Nuclear Energy*, 52: 747–752, 2010. 92, 93, 102, 178

[28] G. G. M. Coppa, V. Giusti, B. Montagnini, and P. Ravetto. On the relation between spherical harmonics and simplified spherical harmonics methods. *Transport Theory and Statistical Physics*, 39:164–191, 2011. 75, 86

[29] R. Dautray, J.L. Lions, and I.N. Sneddon. *Mathematical Analysis and Numerical Methods for Science and Technology: Volume 2 Functional and Variational Methods*. Mathematical Analysis and Numerical Methods for Science and Technology. Springer Berlin Heidelberg, 1999. ISBN 9783540660989. 5, 71

[30] Robert Dautray and Jacques-Louis Lions. *Mathematical Analysis and Numerical Methods for Science and Technology*, volume 6 – Evolution Problems II. Springer, 2000. 8, 9, 19, 20, 21, 23, 34

[31] J.A. Davis. Variational vacuum boundary conditions for a PN approximation. Nov 1965. 43

[32] Timothy A. Davis. Algorithm 832: UMFPACK V4.3—an unsymmetric-pattern multifrontal method. *ACM Transactions On Mathematical Software*, 30(2):196–199, June 2004. ISSN 0098-3500. doi: 10.1145/992200.992206. 123

[33] B. Davison. *Neutron Transport Theory*. Oxford University Press, 1958. 86

[34] Anderson Alvarenga de Moura Meneses, Luca Maria Gambardella, and Roberto Schirru. A new approach for heuristics-guided search in the in-core fuel management optimization. *Progress in Nuclear Energy*, 52(4):339 – 351, 2010. ISSN 0149-1970. doi: http://dx.doi.org/10.1016/j.pnucene.

2009.07.007. URL http://www.sciencedirect.com/science/article/pii/S0149197009001061. 4

[35] L. Demkowicz. *Computing with hp-adaptive Finite Elements: Volume 1. One and Two Dimensional Elliptic and Maxwell Problems*. Chapman & Hall/CRC Applied Mathematics & Nonlinear Science. Taylor & Francis, 2006. ISBN 9781420011685. 51, 120

[36] Jeffery D. Densmore and Ryan G. McClarren. Moment Analysis of Angular Approximation Methods for Time-Dependent Radiation Transport. *Transport Theory and Statistical Physics*, 39:192–233, 2011. 75, 107

[37] Steven Douglass and Farzad Rahnema. Consistent generalized energy condensation theory. *Annals of Nuclear Energy*, 40(1):200 – 214, 2012. ISSN 0306-4549. doi: http://dx.doi.org/10.1016/j.anucene.2011.09.001. URL http://www.sciencedirect.com/science/article/pii/S0306454911003689. 34

[38] Thomas J. Downar. Adaptive nodal transport methods for reactor transient analysis. Technical report, Purdue University, 2005. 74

[39] Pavel Drábek and Jaroslav Milota. *Methods of Nonlinear Analysis*. Birkhäuser, 2007. 26, 27, 60

[40] Michael J. Driscoll and Pavel Hejzlar. Reactor physics challenges in Gen-IV reactor design. *Nuclear Engineering and Technology*, 37(1):1–10, 2005. 51

[41] Jose I. Duo, Yousry Y. Azmy, and Ludmil T. Zikatanov. A posteriori error estimator and {AMR} for discrete ordinates nodal transport methods. *Annals of Nuclear Energy*, 36(3):268 – 273, 2009. ISSN 0306-4549. doi: http://dx.doi.org/10.1016/j.anucene.2008.12.008. URL http://www.sciencedirect.com/science/article/pii/S0306454908003368. {PHYSOR} 2008. 159

[42] Jim E. Morel Edward W. Larsen and John M. McGhee. Asymptotic Derivation of the Multigroup P1 and Simplified PN equations with Anisotropic

Scattering. Technical report, Los Alamos National Laboratory, 1995. 61, 74, 77

[43] Herbert Egger and Matthias Schlottbom. An Lp theory for stationary radiative transfer. *Applicable Analysis*, 93(6):1283–1296, 2014. doi: 10.1080/00036811.2013.826798. 3, 60

[44] P. Solin et al. *Hermes - Higher-Order Modular Finite Element System (User's Guide)*. URL `http://hpfem.org/`. 5

[45] D. Fournier, R. Le Tellier, and C. Suteau. Analysis of an a posteriori error estimator for the transport equation with {SN} and discontinuous galerkin discretizations. *Annals of Nuclear Energy*, 38(2–3):221 – 231, 2011. ISSN 0306-4549. doi: http://dx.doi.org/10.1016/j.anucene.2010.11.006. URL `http://www.sciencedirect.com/science/article/pii/S0306454910003956`. 159

[46] D. Fournier, R. Herbin, and R. L. Tellier. Discontinuous galerkin discretization and *hp*-refinement for the resolution of the neutron transport equation. *SIAM Journal on Scientific Computing*, 35(2):A.936–A.956, 2013. URL `http://search.proquest.com/docview/1322998465?accountid=119841`. 159

[47] Martin Frank. Approximate models of radiative transfer. *Bulletin of the Institute of Mathematics*, 2(2):409–432, 2007. 39

[48] Martin Frank, Axel Klar, Edward W. Larsen, and Shugo Yasuda. Time-dependent simplified PN approximation to the equations of radiative transfer. *Journal of Computational Physics*, 226:2289–2305, 2007. 74

[49] W. Freeden and M. Schreiner. *Spherical Functions of Mathematical Geosciences: A Scalar, Vectorial, and Tensorial Setup*. Advances in Geophysical and Environmental Mechanics and Mathematics. Springer, 2008. ISBN 9783540851127. 46, 48, 85, 89, 90, 161, 162, 163

[50] MA Freitag and A Spence. Convergence theory for inexact inverse iteration applied to the generalised nonsymmetric eigenproblem. *Electronic Transactions on Numerical Analysis*, 28:40–67, 2007. URL `http://opus.bath.ac.uk/172/`. This is the author's final, peer-reviewed version of this document, posted with the publisher's permission. 121

[51] E. M. Gelbard. Application of spherical harmonics methods to reactor problems. Technical Report WAPD-BT-20, Bettis Atomic Power Laboratory, 1960. 52, 73

[52] E. M. Gelbard. Simplified spherical harmonics equations and their use in shielding problems. Technical Report WAPD-T-1182, Bettis Atomic Power Laboratory, 1961. 52, 73

[53] Christophe Geuzaine and Jean-François Remacle. Gmsh: A 3-d finite element mesh generator with built-in pre- and post-processing facilities. *International Journal for Numerical Methods in Engineering*, 79(11):1309–1331, 2009. ISSN 1097-0207. doi: 10.1002/nme.2579. URL `http://dx.doi.org/10.1002/nme.2579`. 150

[54] Thomas Grätsch and Klaus-Jürgen Bathe. A posteriori error estimation techniques in practical finite element analysis. *Computers & structures*, 83 (4):235–265, 2005. 51, 120

[55] Hervé Guillard, Aleš Janka, and Petr Vaněk. Analysis of an algebraic Petrov–Galerkin smoothed aggregation multigrid method. *Applied Numerical Mathematics*, 58(12):1861–1874, 2008. 51

[56] R. Hartmann. Numerical analysis of higher order discontinuous galerkin finite element methods, 2008. URL `http://elib.dlr.de/57074/1/Har08b.pdf`. 69, 70

[57] Vicente Hernandez, Jose E. Roman, and Vicente Vidal. SLEPc: A scalable and flexible toolkit for the solution of eigenvalue problems. *ACM Trans. Math. Software*, 31(3):351–362, 2005. 6, 35, 147

[58] M.E. Hochstenbach and Y. Notay. The jacobi–davidson method. *GAMM-Mitteilungen*, 29(2):368–382, 2006. ISSN 1522-2608. doi: 10.1002/gamm.201490038. 149

[59] Gilles Flamant Hong-Shun Li and Ji-Dong Lu. Mitigation of ray effects in the discrete ordinates method. *Numerical Heat Transfer, Part B: Fundamentals*, 43(12):445–466, 2003. 54

[60] Mohammad Hosseini and Naser Vosoughi. Development of a VVER-1000 core loading pattern optimization program based on perturbation theory. *Annals of Nuclear Energy*, 39(1):35 – 41, 2012. ISSN 0306-4549. doi: http://dx.doi.org/10.1016/j.anucene.2011.08.028. URL `http://www.sciencedirect.com/science/article/pii/S0306454911003665`. 4

[61] Brian Hunter and Zhixiong Guo. Comparison of quadrature schemes in dom for anisotropic scattering radiative transfer analysis. *Numerical Heat Transfer, Part B: Fundamentals*, 63(6):485–507, 2013. doi: 10.1080/10407790.2013.777644. 62

[62] Mathieu Hursin, Shanjie Xiao, and Tatjana Jevremovic. Synergism of the method of characteristic, R-functions and diffusion solution for accurate representation of 3D neutron interactions in research reactors using the AGENT code system. *Annals of Nuclear Energy*, 33:1116–1133, 2006. 37

[63] Alain Hébert. Mixed-dual implementations of the simplified PN method. *Annals of Nuclear Energy*, 37(4):498 – 511, 2010. ISSN 0306-4549. 132, 134, 135

[64] Tudor W. Johnston. General Spherical Harmonic Tensors in the Boltzmann Equation. *Journal of Mathematical Physics*, 7(8):1453–1458, 1966. 90, 92, 93

[65] G. Kanschat, E. Meinköhn, R. Rannacher, and R. Wehrse. *Numerical Methods in Multidimensional Radiative Transfer*. Mathematics and Statistics. Springer London, Limited, 2008. ISBN 9783540853695. 70

[66] D. W. Kelly, J. P. De S. R. Gago, O. C. Zienkiewicz, and I. Babuska. A posteriori error analysis and adaptive processes in the finite element method: Part i—error analysis. *International Journal for Numerical Methods in Engineering*, 19(11):1593–1619, 1983. ISSN 1097-0207. doi: 10.1002/nme.1620191103. 120

[67] R. Kirby. From functional analysis to iterative methods. *SIAM Review*, 52(2):269–293, 2010. doi: 10.1137/070706914. URL `http://dx.doi.org/10.1137/070706914`. 159

[68] W. Kirschenmann, L. Plagne, A. Poncot, and S. Vialle. Parallel SPN on Multi-Core CPUs and Many-Core GPUs. *Transport Theory and Statistical Physics, 39:255-281, 2011*, 39:255–281, 2011. 74

[69] T. Kozlowski and T. J. Downar. Pwr mox/uo2 core transient benchmark. Technical report, 2007. NEA/NSC/DOC(2006)20. 4, 36, 149, 150

[70] G. Hansen H. Park L. Dubcova, P. Solin. Comparison of Multimesh hp-FEM to Interpolation and Projection Methods for Spatial Coupling of Reactor Thermal and Neutron Diffusion Calculations. *J. Comput. Physics*, 230: 1182–1197, 2011. 139, 140, 141

[71] Edward W. Larsen. Asymptotic Diffusion and Simplified PN Approximations for Diffusive and Deep Penetration Problems. Part 1: Theory. *Transport Theory and Statistical Physics*, 39:110–163, 2011. 75

[72] Edward W Larsen, J.E Morel, and Warren F Miller Jr. Asymptotic solutions of numerical transport problems in optically thick, diffusive regimes. *Journal of Computational Physics*, 69(2):283 – 324, 1987. URL `http://www.sciencedirect.com/science/article/pii/0021999187901707`. 63

[73] Edward W. Larsen, Guido Thömmes, Axel Klar, Mohammed Seaïd, and Thomas Götz. Simplified PN Approximations to the Equations of Radiative Heat Transfer and Applications. *Journal of Computational Physics*, 183: 652–675, 2002. 74

[74] D. Lathouwers. Goal-oriented spatial adaptivity for the {SN} equations on unstructured triangular meshes. *Annals of Nuclear Energy*, 38(6):1373 – 1381, 2011. ISSN 0306-4549. doi: http://dx.doi.org/10.1016/j.anucene.2011.01.038. URL `http://www.sciencedirect.com/science/article/pii/S0306454911000594`. 159

[75] R.J. LeVeque. *Finite Volume Methods for Hyperbolic Problems*. Cambridge Texts in Applied Mathematics. Cambridge University Press, 2002. ISBN 9780521009249. 42

[76] Chaung Lin and Shao-Chun Hung. Automatic multi-cycle reload design of pressurized water reactor using particle swarm optimization algorithm and local search. *Annals of Nuclear Energy*, 59(0):255 – 260, 2013. ISSN 0306-4549. doi: http://dx.doi.org/10.1016/j.anucene.2013.04.013. URL `http://www.sciencedirect.com/science/article/pii/S0306454913002156`. 4

[77] Anders Logg and Garth N. Wells. DOLFIN: Automated Finite Element Computing. *ACM Transactions on Mathematical Software*, 37(2), 2010. doi: 10.1145/1731022.1731030. 6, 147

[78] Anders Logg, Kent-Andre Mardal, Garth N. Wells, et al. *Automated Solution of Differential Equations by the Finite Element Method*. Springer, 2012. ISBN 978-3-642-23098-1. doi: 10.1007/978-3-642-23099-8. 6, 68, 147

[79] Thomas A. Manteuffel, Klaus J. Ressel, and Gerhard Starke. A boundary functional for the least-squares finite-element solution of neutron transport problems. *SIAM Journal on Numerical Analysis*, 37(2):556–586, 2000. 23

[80] G Marleau, A Hébert, and R Roy. A user guide for dragon version5. Technical report, Ecole Polytechnique de Montreal, 2014. 37, 127, 146

[81] James C. Maxwell. *A treatise on electricity and magnetism*, volume I. Clarendon Press, Oxford, UK, 1873. 90

[82] Ryan G. McClarren. *Spherical Harmonics Methods for Thermal Radiation Transport*. PhD thesis, The University of Michigan, 2006. URL `http://www.drryanmc.com/papers/2006/McClarren.pdf`. 170

[83] Ryan G. McClarren. Theoretical Aspects of the Simplified Pn Equations. *Transport Theory and Statistical Physics*, 39:73–109, 2011. 74, 86

[84] Ryan G. McClarren and Cory D. Hauck. Robust and accurate filtered spherical harmonics expansions for radiative transfer. *Journal of Computational Physics*, 227:2864–2885, 2008. 49

[85] Ryan G. McClarren, James P. Holloway, and Thomas A. Brunner. On solutions to the Pn equations for thermal radiative transfer. *Journal of Computational Physics*, 229:5597–5614, 2010. 49

[86] Michael F. Modest and Jun Yang. Elliptic PDE formulation and boundary conditions of the spherical harmonics method of arbitrary order for general three-dimensional geometries. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 109(9):1641 – 1666, 2008. URL `http://www.sciencedirect.com/science/article/pii/S0022407307003676`. 85

[87] M. Mokhtar-Kharroubi. *Mathematical topics in neutron transport theory*. World Scientific Pub. Co. Inc., 1998. 28

[88] E. Olbrant, E.W. Larsen, M. Frank, and B. Seibold. Asymptotic derivation and numerical investigation of time-dependent simplified equations. *Journal of Computational Physics*, 238(0):315 – 336, 2013. ISSN 0021-9991. doi: http://dx.doi.org/10.1016/j.jcp.2012.10.055. URL `http://www.sciencedirect.com/science/article/pii/S0021999112006614`. 74

[89] I. Dolezel P. Solin, J. Cerveny. Arbitrary-level hanging nodes and automatic adaptivity in the hp-fem. *Math. Comput. Simul.*, 77:117 – 132, 2008. 5, 112, 119

[90] I. Dolezel P. Solin, K. Segeth. *Higher-Order Finite Element Methods*. Chapman & Hall / CRC Press, 2003. 68, 112, 127, 129

[91] J. Cerveny M. Simko P. Solin, D. Andrs. PDE-Independent Adaptive hp-FEM Based on Hierarchic Extension of Finite Element Spaces. *J. Comput. Appl. Math.*, 233:3086–3094, 2010. 119

[92] L. Dubcova D. Andrs P. Solin, J. Cerveny. Monolithic Discretization of Linear Thermoelasticity Problems via Adaptive Multimesh hp-FEM. *J. Comput. Appl. Math*, 2009. doi: doi10.1016/j.cam.2009.08.092. 5, 109, 113

[93] G. C. Pomraning. Asymptotic and variational derivations of the simplified PN equations. *Annals of Nuclear Energy*, 20:623–637, 1993. 74

[94] Jean C. Ragusa. Application of h-, p-, and hp-mesh adaptation techniques to the SP3 equations. *Transport Theory and Statistical Physics*, 39:234–254, 2011. 74

[95] Jean C Ragusa and Yaqi Wang. A two-mesh adaptive mesh refinement technique for SN neutral-particle transport using a higher-order DGFEM. *Journal of computational and applied mathematics*, 233(12):3178–3188, 2010. 159

[96] A. Ralston and P. Rabinowitz. *A First Course in Numerical Analysis: Second Edition*. Dover Books on Mathematics. Dover Publications, 2012. ISBN 9780486140292. 64

[97] Paul Reuss. *Neutron Physics*. EDP Sciences, 2008. 37, 39

[98] E. Romero and J. E. Roman. A parallel implementation of Davidson methods for large-scale eigenvalue problems in SLEPc. *ACM Trans. Math. Software*, 40(2):13:1–13:29, 2014. 149

[99] G.Ya. Rumyantsev. Boundary conditions in the spherical-harmonic method. *Journal of Nuclear Energy. Parts A/B. Reactor Science and Technology*, 14(1–4):215 – 223, 1961. ISSN 0368-3230. doi: http://dx.doi.org/10.1016/0368-3230(61)90125-2. URL `http://www.sciencedirect.com/science/article/pii/0368323061901252`. 43, 86

[100] Richard Sanchez. Treatment of boundary conditions in trajectory-based deterministic transport methods. *Nuclear Science and Engineering*, 140: 23–50, 2002. 12

[101] Richard Sanchez. The Criticality Eigenvalue Problem for the Transport Operator with General Boundary Conditions. *Transport Theory and Statistical Physics*, 35(5):159–185, 2006. 12, 20, 27, 28

[102] Richard Sanchez. Prospects in Deterministic Three-dimensional Whole-core Transport Calculations. *Nuclear Engineering and Technology*, 44(2): 113–150, 2012. 3, 37

[103] Richard Sanchez. On P_N Interface and Boundary Conditions. *Nuclear Science and Engineering*, 177:19–34, 2014. 43

[104] S. Santandrea, R. Sanchez, and P. Mosca. A linear surface characteristics approximation for neutron transport in unstructured meshes. *Nuclear Science and Engineering*, 160:23–40, 2008. 37

[105] D. S. Selengut. A new form of the P3 approximation. *Transactions of American Nuclear Society*, 13:625–626, 1970. 107, 159

[106] Gerard L.G. Sleijpen and Henk A. van der Vorst. An overview of approaches for the stable computation of hybrid bicg methods. *Applied Numerical Mathematics*, 19(3):235 – 254, 1995. ISSN 0168-9274. Iterative Methods for Linear Equations. 149

[107] P. Solin. *Partial Differential Equations and the Finite Element Method*. J. Wiley & Sons, 2005. 23, 69, 111, 120

[108] Pavel Solin and Stefano Giani. An iterative adaptive finite element method for elliptic eigenvalue problems. *Journal of Computational and Applied Mathematics*, 236(18):4582–4599, 2012. 51

[109] Weston M. Stacey. *Nuclear Reactor Physics*. John Wiley & Sons, Inc., New York, 2nd edition, 2007. 23, 39, 51

[110] R. J. J. Stamm'ler and M. J. Abbate. *Methods of Steady-State Reactor Physics in Nuclear Design*. Academic Press, 1983. 39

[111] Jarred Tanner. Optimal filter and mollifier for piecewise smooth spectral data. *Mathematics of Computation*, 75(254):767–790, 2006. 49

[112] P. Vaněk. Acceleration of convergence of a two-level algorithm by smoothing transfer operator. *Appl. Math.*, (37), 1992. 179

[113] P. Vaněk. Fast multigrid solver. *Applications of Mathematics*, 40(1), 1995. 179

[114] P. Vaněk, J. Mandel, and M. Brezina. Algebraic multigrid by smoothed aggregation for second and fourth order elliptic problems. *Computing*, 56, 1996. 179

[115] P. Vaněk, M. Brezina, and R. Tezaur. Two-grid method for linear elasticity on unstructured meshes. *SIAM J. Sci Comput.*, (21), 1999. 185, 186

[116] P. Vaněk, M. Brezina, and J. Mandel. Convergence of algebraic multigrid based on smoothed aggregations. *Numer. Math.*, 88(3), 2001. 179, 180, 186

[117] V.S. Vladimirov and Atomic Energy of Canada Limited. *Mathematical Problems in the One-velocity Theory of Particle Transport*. Atomic Energy of Canada Limited, 1963. 17, 22

[118] Yaqi Wang and Jean C Ragusa. On the convergence of DGFEM applied to the discrete ordinates transport equation for structured and unstructured triangular meshes. *Nuclear science and engineering*, 163(1):56–72, 2009. 160

[119] Yaqi Wang and Jean C. Ragusa. Standard and goal-oriented adaptive mesh refinement applied to radiation transport on 2d unstructured triangular meshes. *Journal of Computational Physics*, 230(3):763 – 788, 2011. ISSN 0021-9991. doi: http://dx.doi.org/10.1016/j.jcp.2010.10.018. URL `http://www.sciencedirect.com/science/article/pii/S0021999110005747`. 159

[120] Y. Watnabe and Maynard C. W. The discrete cones method in two-dimensional neutron transport computation. Technical report, Fusion Technology Institute, University of Wisconsin, 1984. 145

[121] E.W. Weisstein. *CRC Concise Encyclopedia of Mathematics, Second Edition*. Taylor & Francis, 2002. ISBN 9781420035223. URL `http://books.google.cz/books?id=aFDWuZZslUUC`. 65

[122] G.J. Wu and R. Roy. A new characteristics algorithm for 3D transport calculations. *Annals of Nuclear Energy*, 30:1–16, 2003. 37

[123] V. G. Zimin. Personal communication. National Research Nuclear University MEPhI, Moscow, 2011. 4

Index

INDEX

INDEX

List of publications

Research reports

1. M. Hanuš., R. Kužel. SMV13 Souhrná výzkumná zpráva. Detailní analýza implementačních možností mnohagrupového SP3 modelu pro nodální kód Hanka. 2013.

2. M. Hanuš., R. Kužel. SMV13 Souhrná výzkumná zpráva. Dvou dimenzionální rekonstrukce neutronového toku na hexagonální oblasti z nodálních výpočtů. 2013.

3. R. Čada, M. Hanuš., R. Kužel, J. Prehradný. SMV12 Souhrná výzkumná zpráva. 2012.

Articles in journals

1. M. Brandner, M. Hanuš, and R. Kužel. Nodal methods for a two-dimensional static multigroup diffusion calculation of nuclear reactors with hexagonal assemblies. *Journal of Interdisciplinary Mathematics*, 12(2):203-224, 2009. ISSN: 0972-0502.

2. Fraňková, P., Hanuš, M., Kopincová, H., Kužel, R., Marek, I., Pultarová, I., Vaněk, P., Vastl, Z.: *Convergence theory for the exact interpolation scheme with approximation vector used as the first column of the prolongator: the partial eigenvalue problem.*
   Submitted to Numerische Mathematik.

Contributions to conference proceedings

1. M. Hanuš. A new perspective on some approximations used in neutron transport modeling. In *Proceedings of the Seminar Programs and Algorithms of Numerical Mathematics 16*, pp. 81-87. Prague: Institute of Mathematics, Academy of Sciences of the Czech Rep., 2013.
   ISBN 978-80-85823-62-2.

2. P. Fraňková, M. Hanuš, H. Kopincová, R. Kužel, P. Vaněk, Z. Vastl. A short philosophical note on the origin of smoothed aggregations. In *Proceedings of the International Conference Applications of Mathematics 2013*, pp. 67-76. Prague: Institute of Mathematics, Academy of Sciences of the Czech Rep., 2013.
   ISBN 978-80-85823-61-5.

3. M. Hanuš. On the Development of a New Optimization Code for Nuclear Power Plants. In *Proc. of the International Conference on Mathematics in Engineering & Business Management*, pp. 168-171. Chennai: Stella Maris College, India, 2012.
   ISBN: 978-81-8286-015-5.

4. M. Hanuš, M. Kadlecová. Numerické metody vyššího řádu pro řešení transportních úloh. In *Proc. of the Seminar on numerical analysis.* pp. 39-42. Ostrava: Ústav geoniky AV ČR, 2011.
   ISBN: 978-80-86407-19-7.

5. M. Hanuš. Vývoj optimalizačního kódu pro jaderné elektrárny.
   In *Jaderná energetika, transmutační a vodíkové technologie v pracích mladé generace - 2011, Sborník referátů ze semináře*, pp. 176-179. Praha, 2012.
   ISBN: 978-80-02-02360-9

6. M. Hanuš. Moderní numerické metody pro neutroniku a sdružené úlohy.
   In *Jaderná energetika, transmutační a vodíkové technologie v pracích mladé generace - 2010, Sborník referátů ze semináře*, pp. 115-120. Praha, 2011.
   ISBN: 978-80-02-02288-6.

7. M. Hanuš. Modelování neutronových toků.
   In *Jaderná energetika, transmutační a vodíkové technologie v pracích mladé generace - 2009, Sborník referátů ze semináře*, pp. 25-33. Praha, 2010.
   ISBN: 978-80-02-02209-1.

8. T. Berka, M. Brandner, M. Hanuš, R. Kužel and A. Matas. A 3D model of neutron flux in nuclear reactors with hex-shaped assemblies. In *Proceedings of the Seminar Programs and Algorithms of Numerical Mathematics 14*, pp. 83-90. Prague: Institute of Mathematics, Academy of Sciences of the Czech Rep., 2008. ISBN: 978-80-85823-55-4.

Monograph

M. Hanuš. *Mathematical Modeling of Neutron Transport: Theoretical and computational point of view.* 1st Ed. Saarbrücken: LAP LAMBERT Academic Publishing GmbH & Co. KG, 2011, 132 p.
ISBN: 978-3-8443-0121-2.